UNIVERSITY OF LONDON

Coordination and Collaboration Within and Across Organizations: The Role of Experience and Knowledge on Innovation

A DISSERTATION

submitted in partial fulfilment of the requirements for the degree

DOCTOR OF PHILOSOPHY

Field of Strategy and Entrepreneurship

London Business School

By

M. Onal Vural

London, U.K.

May, 2011

ABSTRACT

This thesis represents a conceptual and empirical investigation of how coordination problems are resolved in scientific teams and across organizations working in knowledge intensive environments. It takes a distinct view of coordination by bridging sociology of science, organizational learning and capability development perspectives. In the developed theoretical model, effective coordination mechanisms are associated with experience in joint work and learning as well as design choices in terms of complexity and knowledge relevance.

In testing the theoretical constructs, the thesis uses a four study format; the first study analyzes resolution of coordination problems stemming from diversity at the team level. Drawing on the sociology of science and team work literatures, this study distinguishes between effects of task specific and domain specific expertise in resolving coordination problems in a context where familiarity with previous knowledge becomes detrimental. The aim of the second and third studies is to focus on the resolution of inter-organizational coordination problems stemming from complexity in project design and knowledge relevance in the distributed work environment. Through empirical analysis of clinical trials, these studies show that prior experience in collaboration fosters resolution of coordination problems. This effect is enhanced when the prior experience stems from the same knowledge base as the focal project and when prior experience is gained from working together in complex projects.

The final study draws on 30 years of clinical trials in a well codified knowledge setting of Type II Diabetes trials. Focusing on the effect of complexity on delegation of distributed work, the study analyzes how collocation enhances resolution of coordination problems in a setting

where work distribution is becoming increasingly international. On the overall, this thesis focuses on the resolution of inter-personal and inter-organizational coordination issues in joint work at two levels of analysis and outlines a number of theoretical contributions, directions for future research, and implications for practice.

Acknowledgements

I am indebted to many people for guiding me and supporting me during my doctoral studies. First of all, I would especially like to thank my supervisor Professor Gerard George for his continuous support during my studies. Gerry is truly a great mentor and a source of inspiration. I thank him for providing me with guidance in my work and for encouraging me in my path towards academic development. He has had a significant impact in both shaping my research and shaping my approach to conducting research.

I am also very grateful to Professor Phanish Puranam and Professor Madan Pilllutla for acting as examiners on my dissertation proposal committee. Their insightful comments were very valuable in the development of my work. Also, Professor Puranam supplied his valuable comments and guidance as I progressed towards completion.

I would like to thank the Strategy and Entrepreneurship group at London Business School for supporting me and making me feel at home after my transition from Entrepreneurship group into the Strategy area. Faculty in the Strategy and Entrepreneurship group have extended their valuable comments on my work throughout the past five years. I would like to thank Dr. Lourdes Sosa for being on my examination committee during my second year and for her insightful comments on my work all throughout the program. I would also like to thank Dr. Brandon Lee, Dr. Kevin Boudreau, Dr. Donal Crilly and Dr. Markus Reitzig for their comments on earlier versions of my work.

In addition, I would like extend my gratitude to my viva examiners, Professor Erkko Autio and Dr. Freek Vermeulen, who have kindly accepted to act as examiners on my examination committee.

My colleagues in the Strategy group provided me with some of the best advice and insights that I have received both on my work and on the PhD process during my time at London Business School. I would like to extend my thanks to Dr. Reddi Kotha, Dr. Kannan Srikanth, Dr. Felipe Monteiro, Dr. Susan Hill, Dr. Mihaela Stan, Dr. Kai-yu Hsieh, Marko Coh, Eucman Lee, Ramon Leucona, Maya Cara, Marlo Goetting and Jennifer Tae.

In addition my dear friends from the Finance department were there to support me throughout this journey as well, a special thanks goes to Marc Gabarro, Ramin Baghai, Dr. Oguzhan Karakas and Dr. Jungsuk Han.

Most important of all, I would like to thank my brother Yenal Vural and my parents Halil and Bedia Vural for their unconditional support on my endeavours. It would have been impossible for me to be who I am right now without their encouragement and support.

TABLE OF CONTENTS

LIST OF TABLES

CHAPTER 1

Introduction

1.1 Background

The focus of this thesis is on the design, organization and execution of collaborative activity where parties from diverse backgrounds with different levels of experience work in tandem. Collaborations have become a norm in knowledge intensive domains where management of new knowledge creation and innovation is of essential value. Such settings include team work in and across organizations in technology intensive areas as well as inter-organizational work in distributed work environments (Ahuja, 2000; Kiesler *et al.*, 2002; Powell *et al.*, 1999; Powell *et al.*, 1996; Rothaermel *et al.*, 2007; Zucker *et al.*, 1996).

The creation of new knowledge is deemed one of the most important factors for the success of organizations (Grant, 1996). Knowledge is created as parties combine and exchange resources and information. Obtaining resources and developing knowledge through interactions with one another is one of the ways in which individuals create knowledge (Polanyi, 1967). Research on innovation and scientific discoveries has shown that, in fact, the locus of knowledge creation has shifted from individuals to teams and to networks of knowledge workers thus bringing about a greater necessity to understand the build up of effective collaboration routines and design of team structures in innovative domains (Ahuja, 2000; Wuchty *et al.*, 2007).

At the same time, a growing number of projects require the coordination of activities across organizational boundaries due to increasingly distributed nature of work in knowledge intensive industries on a global scale (Kiesler *et al.*, 2002; Kraut *et al.*, 2002; Srikanth, 2010; Srikanth *et al.*, 2009). This increasing preponderance of collaborative activity brings to the fore the issues regarding the design of collaborative work in innovative settings and the coordination

of distributed work given inter-organizational differences in organizational attributes, project attributes and collaborating parties' prior experiences.

As collaborations come to the fore of organizing at both inter-organizational and team level innovative activity, a greater need to understanding the design and execution of collaborations arises. An important feature of collaborative activity is the involvement of multiple actors whose performance is interdependent on the work of one another. Hence coordination becomes an important aspect of collaborations where actions of interdependent actors need to be aligned through reciprocal predictability of action (Camerer *et al.*, 2003; Heath *et al.*, 2000; Simon, 1997; Srikanth, 2010).

Although scholars have taken various lenses to study such inter-organizational and inter-unit collaborations, a unification of approaches in the design and effective management of collaborative work given prior experiences and characteristics of collaborating parties has been lacking.

1.2 Research Questions

In this dissertation I address the central question: how does prior experience stemming from different types of joint work mitigate coordination concerns in collaborative projects? In addressing this question, I shift the focus away from individual collaborator characteristics to design and management of collaborative work. The unit of analysis is a collaborative project; in the case of the first study the unit of analysis becomes the team work in inventions in human embryonic stem cell (hESC) research domain, in the three following studies using clinical trial databases the unit of analysis is the clinical trial project.

In the first study, I begin by focusing on the organization of collaborative projects in scientific teams. By analyzing the structure of scientific teams working on inventions in a new

and highly dynamic area, human embryonic stem cell research, I ask two related questions: a) how can members of a scientific team effectively coordinate their action while bridging institutional, geographic and knowledge gaps and b) to what extent individual and joint work experiences help mitigate coordination problems when such experience stems from within vs. outside the focal knowledge domain?

In the second study, I analyze the design of distributed work projects and how complexities in design affect the performance of distributed work. I use clinical trials as a setting and explore the complexities associated with the design of a distributed work project in clinical trials. By establishing appropriate measures of complexity and the corresponding effects on project performance I set the stage for my third study.

Chapter 3, sets the stage for Chapter 4 which focuses on not only the dimensions of complexity in designing distributed work projects but also empirically investigates the effect of organizational and inter-organizational experiences on project performance. The question I ask in this study is: how do the relevance of knowledge and intensity of task complexity affect the coordination benefits that stem from having experience in inter-organizational joint work? I analyze how project characteristics affect learning and how build up of relational capital from previous projects impacts the performance of focal project.

In my fourth study, I focus particularly on the location choices of organizations when delegating tasks in a globally distributed work setting. I ask the questions: what are the effects of internationalization and work distribution on the management of distributed work projects in a well codified knowledge setting, and to what extent does co-location matter in a globally distributed work project. I analyze how different coordination mechanisms, name collocation and planning, affect the performance internationally distributed work in clinical trials.

Studying the differences between coordination enhancing experience effects stemming from knowledge specific and complex tasks in collaborative projects at team and inter-organizational levels is important for the following reasons. First, it will give us a better understanding of what experience in joint work does for the effective management of collaborative projects. This topic is central to many coordination studies, yet there is little empirical evidence as to how different types of experience mitigate coordination concerns in collaborative work. If design of the collaborative projects is important (and the previous research suggests that it is), then understanding design choices related to dimensions of complexity in collaborative projects is essential. Second a greater understanding of these relationships across design and execution of collaborative projects should help us build more prescriptively valid theories. Even though more accurate assumptions do not always lead to better predictions, a greater understanding could point to important boundary conditions of studied relationships. Third, a more profound understanding of these relationships should increase the descriptive validity of theories. Such understanding may be crucial when these theories are used to make normative statements (Vanneste, 2009).

1.3 Literature Review

In this section, I provide a review of the most relevant literature on the organization of collaborative activity and the related coordination concerns. I focus on three levels of collaborative activity; a) design of collaborative work at the team level, b) design and coordination of distributed work at the inter-organizational level, and c) coordination of distributed work across boundaries and the choice of locations.

1.3.1 Collaboration and Knowledge Creation in Science

The literature on collaboration and knowledge creation in science includes works from scholars who have looked at joint work in innovation and how the quality of the collaboration processes affects the success of collaborative team work (Hoegl *et al.*, 2001), as well as the processes associated with the generation of knowledge through an accumulative logic (Murray *et al.*, 2007a; Murray *et al.*, 2007b). Moreover, scholars have observed how the diffusion of scientific advancements takes place and how norms of disclosure affect this process (Sorenson *et al.*, 2004).

Research on new knowledge creation particularly focuses on recombination processes, depicting new knowledge creation to occur through novel combinations of existing knowledge (Fleming, 2001, 2002; Fleming *et al.*, 2004). In this regard, how network positions of organizations influence the organizations' brokerage roles in the knowledge creation processes has been analyzed (Hargadon, 2003; Hargadon *et al.*, 1997).

Some of the topics investigated in collaborative knowledge creation efforts include; issues on definition and measurement of collaboration (Katz *et al.*, 1997), learning from collaborations (Reagans *et al.*, 2005) and collaboration across institutions (Cummings *et al.*, 2007). Moreover, researchers have analyzed at team level, the antecedents of successful scientific collaborations crossing disciplinary boundaries (Cummings *et al.*, 2005).

Previous studies have shown a positive link between a team-based organization and innovative performance (Gupta *et al.*, 1996), but with limited specificity of its impact on dimensions of success in innovative projects (Hoegl *et al.*, 2001). The ability of the individuals taking part in a team to collaborate and the interdependence of team members' efforts, as well as the diversity and depth of knowledge they bring to the project determine the project outcomes

(Gibson *et al.*, 2003; Taylor *et al.*, 2006). A major problem with isolating the drivers of success in team work stems from the fact that team success depends on collaborative work of individuals, who bring different sets of knowledge and experience.

In studying the benefits of collaborative work, scholars have analyzed how different types of alliance experiences; especially partner specific and general alliance experience relate to the performance of joint projects (Hoang *et al.*, 2005) as well as how team level learning takes place when teams adopt disruptive technologies (Edmondson *et al.*, 2001) and how learning and improvisation processes occur in groups (Miner *et al.*, 2001). Notwithstanding these rich studies, there is limited research on collaborative projects in scientific research teams with respect to the development of collaborative capabilities (Taylor *et al.*, 2006; Tzabbar, 2009) and overcoming coordination problems (Kiesler *et al.*, 2002).

1.3.2 Inter-organizational Collaboration and Coordination

The scope of scholarly work in inter-organizational relations and coordination of joint work has particularly been analyzed in the alliance literature (Kale *et al.*, 2007; Schreiner *et al.*, 2009). This work takes a knowledge based view (Grant, 1996) and includes capability development perspectives (Hoang *et al.*, 2005; Hoegl *et al.*, 2004; Kale *et al.*, 2002; Kale *et al.*, 2007) in answering how organizations effectively coordinate inter-organizational work.

Scholars have suggested that organizations can develop an ability to effectively manage inter-organizational relationships by having greater experience in managing such relationships (Anand *et al.*, 2000; Hoang *et al.*, 2005; Schreiner *et al.*, 2009; Zollo *et al.*, 2002), and they have found empirical support for this claim. Moreover, coordination by planning and design of joint work has been shown as an effective way of managing inter-organizational activity (Hoang *et al.*, 2005; Kale *et al.*, 2002; Schreiner *et al.*, 2009). At the same time, recent work has shown that

organizations can implement practices to capture, codify and internalize relevant know-how that enhances the organizations' ability to effectively manage inter-organizational joint work (Kale *et al.*, 2007).

Research in inter-organizational joint work points to factors affecting the success of such collaborations based on the organizations' ability to select partners that have the right complementarities and fit (Dyer *et al.*, 1998; Hitt *et al.*, 2000) and the right organizational structures (Gulati *et al.*, 1998; Hennart *et al.*, 2005; Schreiner *et al.*, 2009). At the same time the success of collaborative work is also shown to be related to the ability of organizations to manage such work in terms of task coordination, conflict resolution and knowledge sharing (Doz, 1996; Kumar *et al.*, 1998; Madhok *et al.*, 1998; Van de Ven *et al.*, 1992).

Inter-organizational collaborations face coordination challenges that stem from interdependence between partners (Gerwin, 2004; Gulati *et al.*, 1998; Litwak *et al.*, 1962) that can partly be addressed at partner selection and design stages (Gulati *et al.*, 1998). However, not all coordination challenges can be solved before hand, therefore mechanisms that help build the ability of organizations to effectively manage joint work are of great importance.

The importance of coordination in collaborations is paramount given the fact that partners need to specify roles and responsibilities based on task characteristics and adapt to changes (Gerwin, 2004; Gulati *et al.*, 1998; Lawrence *et al.*, 1967; Litwak *et al.*, 1962; Schreiner *et al.*, 2009). This means that the ability to coordinate stems from knowledge and skills to identify the interdependence and manage it as the need for the adaptation arises. Although routine tasks may be specified easily, more complex and non-routine tasks require higher levels coordination ability. Even though previous research has pointed to the importance of prior experience in inter-organizational relations as a way to develop effective coordination mechanisms (Anand *et al.*,

2000; Hoang *et al.*, 2005; Zollo *et al.*, 2002), little exploration has been done in terms how complexity and nature of knowledge in collaborative projects affect the development of such capabilities.

1.3.3 Inter-organizational Collaboration, Coordination and Globally Distributed Work

International business literature focuses densely on the issues of knowledge transfer while paying little attention to coordination of distributed work across boundaries (Kumar *et al.*, 2008; Srikanth, 2010). Prior work in this area has been analyzing three approaches to coordinating internationally distributed work: a) coordination through planning, b) coordination through feedback and c) tacit coordination mechanisms.

Coordination through planning mainly deals with task decomposition and modularization in order to avoid the necessity of on-going communication for coordinating action (Galbraith, 1977; March *et al.*, 1958; Tushman *et al.*, 1978). This approach, although feasible in large scale projects where modularization is possible and maintainers with profound knowledge of the system are readily available, is not an option in smaller scale projects where upfront investments in modularization may not be viable (Srikanth, 2010). Therefore organizations may fall back on the alternative option of coordinating through feedback.

Coordination through feedback which refers to on-going communication to overcome conflicts that stem from interdependencies (Hinds *et al.*, 2005) has been associated with collocation as a way to enable communication. The research in this area has focused on the benefits of collocation in terms of sharing a social context and being able to have face-to-face communication (Kiesler *et al.*, 2002). The importance of having shared artefacts, shared conventions and being aware of each others' work has also been emphasized as essential benefits of proximity in improving coordination (Cramton, 2001; Kraut *et al.*, 2002; Olson *et al.*, 2002).

Spatial proximity is also shown to increase spontaneous communication and improve the ease with which knowledge is shared (Kraut *et al.*, 2002). The essence of this coordination mechanism is similar to what is achieved by common ground through mutual knowledge of role responsibilities (Bechky, 2006; Faraj *et al.*, 2006), the environment (Bechky, 2006; Clark, 1996) or background scientific knowledge (Puranam *et al.*, 2009).

Furthermore, a growing line of research deals with tacit coordination mechanisms in off-shore development projects that deal with coordinating action through procedural, contextual and interpersonal common ground. This research deals with tacit mechanisms that build up on the above mentioned common ground schemes to overcome coordination problems through neither planning, nor feedback but by tacit coordination mechanisms (Orlikowski, 2002; Srikanth, 2010).

Although, prior research has focused on various coordination mechanisms to effectively manage internationally distributed work projects, no empirical research has thus far focused on the co-existence of and complementarities between coordination by planning and coordination by feedback mechanisms.

1.3.4 Conclusions from Literature Review

A number of conclusions can be drawn from this literature review. First, although scholars have analyzed coordination problems in teams by looking at team structure and affiliations, no work has dealt with the interaction between prior experience stemming from within and outside the knowledge domain of focal projects and how such experiences affect team level coordination in innovative work.

Second, the development of an ability to effectively manage and coordinate inter-organizational work has been connected to prior experiences (Hoang *et al.*, 2005), however, how

knowledge specificity and complexity of collaborative projects influence the importance of such experience has not been studied.

Third, effective management of internationally distributed work projects as a relatively new area had no empirical research focusing on the extent to which collocation (coordination by feedback) can be influential in a setting where coordination by planning is deemed possible and effective.

1.4 Structure of Dissertation

Because of the theoretical importance of coordination mechanisms in collaborative projects and because of limited evidence on experience based coordination mechanisms, I will focus in this dissertation on the interaction between different types of experience in mitigating coordination concerns in collaborative projects. Further, and as a related line of work, I also analyze the impact of collocation as a coordination mechanism in internationally distributed work. I have managed to find two settings in which I can observe the elements of experience and how they pertain to resolving coordination problems both in teams and in distributed work projects across organizations. I have conducted interviews with prominent figures in the two settings to get a better sense of the phenomena and to direct my theoretical and empirical approach. Although I do not use interview data as a primary data source, having conducted interviews helped me gain a better understanding of the settings that I use in my dissertation. Details regarding the interviews are included in Appendix-A.

I use both interviews and archival data for my studies, in particular I use patent and publication data for my first study and large scale archival databases clinical trials supplemented with indexes of clinical research organizations for the second, third and fourth studies. Below, I give a brief overview of each chapter.

In Chapter 2, I explore the extent to which prior experience in working together helps mitigate coordination concerns in a dynamic scientific setting. In particular, I focus on inventions in human embryonic stem cell research and analyze the effect of collaborative experiences of team members within and outside this setting on the effective coordination of scientific team work.

In Chapters 3 and 4, I explore first the complexities in the design of collaborative work projects in clinical trials and then analyze how the involvement of such complexities in the design of projects affects project performance. Second, by constructing experience measures that interact with the specific knowledge and complexity measures pertaining to the design of distributed work projects, I analyze how different types of experience affect the performance of projects given project attributes.

In Chapter 5, I study how collocation as a choice of coordination mechanism in internationally distributed work influences the performance of projects in an area where effective modularization and coordination by plan mechanisms would be deemed possible.

In Chapter 6, I provide conclusions, offer managerial and theoretical implications and highlight potential threads for future research.

CHAPTER 2

Collaboration and Coordination in Scientific Teams

2.1 Introduction

Scientists aspire for breakthrough inventions that will reshape their discipline. An emerging debate to explain why some breakthrough inventions emerge is the role of collaboration and the influence of diverse scientific backgrounds. Collaboration fosters integration of skills, ideas and experiences across individuals; it also helps develop new insights through the recombination of relevant knowledge across sub-fields. Although scholars have analyzed collaboration in science, the analysis has primarily focused on the issues of teamwork dealing with the choice of collaboration partners and the trend toward joint research (Guimera *et al.*, 2005). This line of thought has established that there is a trend towards cross-institutional collaborations (Jones *et al.*, 2008) as well as difficulties associated with such collaborations (Cummings *et al.*, 2005, 2007).

Previous studies have shown a positive link between a team-based organization and innovative performance (Gupta *et al.*, 1996), but with limited specificity of its impact on dimensions of success in innovative projects (Hoegl *et al.*, 2001). The ability of the individuals taking part in a team to collaborate and the interdependencies among the team's members, as well as the diversity and depth of knowledge they bring to the project determine the project outcomes (Taylor *et al.*, 2006). A major problem with isolating the drivers of success in teamwork stems from the fact that team success depends on collaborative work of individuals, who bring different sets of knowledge and experience.

In studying the benefits of collaborative work, scholars have analyzed how different types of alliance experiences; especially partner specific and general alliance experience relate to

the performance of joint projects (Hoang *et al.*, 2005) as well as how team level learning takes place when teams adopt disruptive technologies (Edmondson *et al.*, 2001) and how learning and improvisation processes occur in groups (Miner *et al.*, 2001). At the firm level, the benefits of collaboration include enhanced firm connectivity to its environment which stimulates innovative processes (Powell *et al.*, 1996). Another stream has focused on the strategic development of collaborative capabilities that describe how firms capture, disseminate and manage relationships across organizational boundaries and how the existence of such capabilities enhances performance outcomes (Dyer *et al.*, 1998; Kale *et al.*, 2002). Notwithstanding these rich studies, there is limited research on collaborative projects in scientific research teams with respect to the development of collaborative capabilities (Taylor *et al.*, 2006; Tzabbar, 2009) and overcoming coordination problems (Kiesler *et al.*, 2002) to invent.

This study aims to fill this gap by studying inventions and patenting activity in human embryonic stem cells (hESC), an area of research with the potential to solve numerous debilitating ailments such as Diabetes, Alzheimer's and Parkinson's disease (McCormick *et al.*, 2009; Scott *et al.*, 2009). Stem cell research is relevant as a context for this study because of its emergence as a technical field of study over the past two decades. Importantly, the cell lines themselves are fickle material that requires significant tacit knowledge on handling and propagation, which makes experience more relevant for success in discovery.

Two factors of theoretical and practical interest come into play when analyzing teamwork success in innovative projects: (1) collaboration benefits arise from the enhanced ability to work together effectively and develop routines and processes, possibly using their previous experiences with similar tasks as a lens; and (2) coordination costs that emerge as the teams seek to bridge the institutional, geographic and cultural gaps among its members. Collaboration is

thus a two-sided sword in the context of radical, new science as it brings both benefits *and* costs. Successful collaborations occur when the benefits from collaboration outweigh the costs of coordination that team members face. Hence, in this study, I systematically examine the effects of prior collaborative experience and the recombination ability among team members on the scientific impact of inventions in radical, new science.

The research setting offers an excellent opportunity to study factors affecting collaboration in geographically and institutionally separated scientific teams and the impact of their invention outcomes. My data cover 12 years of patenting activity since the discovery of human embryonic stem cells (hESC) and gives a comprehensive picture of collaborations in this field. Due to the recent emergence of the hESC field, the work in this area is concentrated among 648 scientists and dominated by collaborative work that I am able to observe in our data. From this data, I draw a range of different variables to assess the effects of experience with a unique resource as well as experience in joint work within and outside the field, while controlling for team-level diversity in terms of scientific and departmental structure of the team.

2.2 Theory and Hypotheses

2.2.1 Generalized Experience (Experience Outside hESC Domain)

Experience accrues two major benefits for a team. First, it provides the benefit of resource sharing, allowing researchers to combine knowledge, skills, and access to physical assets (e.g. labs). Experience helps in developing new routines that facilitate coordination and solve the mutual knowledge and task allocation problems. Improved knowledge codification and enhanced cooperation over collaborative activities results in more effective collaboration. Previous research has shown that firms with prior alliance experience perform better than those without such experience (Hoang *et al.*, 2005). Second, it helps to develop critical information

about who knows what and who knows how that aids in effective communication and task allocation (Cummings *et al.*, 2005, 2007, 2008; Owen-Smith *et al.*, 2003). Teams with individuals who have prior joint work experience are therefore more efficient in carrying out research then teams that lack such experiences (Katz, 1982).

Experience also allows teams to adopt practices that increase their ability to coordinate their activities (Bercovitz *et al.*, 2011), as well as helping establish expectations about acceptable levels of performance. However, the effectiveness of this adoption may work against the team when a change in context requires novel ways of approaching and solving problems as in the case of work involving human embryonic stem cell lines. Previous work routines may no longer be applicable to the changing context of research, and become detrimental to performance of the project. Radical discoveries could also be disruptive to existing routines for scientific production, i.e., the task-oriented day-do-day processes, and reduce the relevance of prior generalized experience for the new context. For example, how cell culture and propagation is carried out substantively differs between embryonic stem cells and other adult mature cells (Jain *et al.*, 2007). Further, in radical discoveries, prior joint experience could hinder experimentation and focus team efforts along established lines of inquiry that apply to mature fields. Taken together, I expect that teams with generalized joint work experience may have a higher tendency to adopt tried-and-tested means to solve novel problems, thereby negatively affecting the quality of the invention.

> *Hypothesis 1: The effect of generalized experience (prior joint work experience outside the hESC domain) of team members on the scientific impact of the focal patent is negative.*

2.2.2 Scientific Proximity

Scholars have posited that new knowledge is created by unique recombination of existing knowledge repositories (Basalla, 1988; Henderson *et al.*, 1994; Schumpeter, 1939). Though

inventors can possibly combine any prevalent technological components, what actually gets combined is constrained by the localness of their search and the social construction on what components can be gainfully combined. Scientific distance can be thought of as a sub-measure of cognitive distance relating to science (Nooteboom, 2000a, 2000b; Nooteboom *et al.*, 2007), which determines the extent to which scientists perceive the world differently from one another based on the development of their cognition stemming from their prior scientific work. In this respect, scientific proximity decreases the stock of opportunities to which scientists have access in their joint work since scientifically proximate individuals can only perceive a narrow spectrum of the paths available (Fleming, 2001; Fleming *et al.*, 2004), that is, teams that recombine ideas from proximate technological niches are likely to have lower scientific impact on their field (George *et al.*, 2008).

In teams with higher scientific distance, the difference in scientific expertise of scientists across multiple topics allow for a variety in problem solving approaches which increases the likelihood that solutions can be found for important technological bottlenecks. Recombination can also enhance the impact of the innovation on the technology domain itself. Indeed, it has been argued that breakthroughs result from recombining non-obvious technology components (Basalla, 1988). Hence, when the scientific proximity is low in a team, the inventing team could combine new knowledge with its existing knowledge to yield radical innovations (Ahuja *et al.*, 2001; Katila *et al.*, 2002) that can potentially influence both domains (Ethiraj *et al.*, 2004).

Though searching widely for technology solutions has positive implications in terms of the space for recombination and the consequent impact, the ability to effectively recombine this knowledge depends on the team's ability to coordinate their activities. As the scientific distance increases across inventors, a scientific language problem arises such that wider gaps lead to

increased difficulty in communicating ideas across the team (Katz, 1982). Moreover, absorptive capacity of scientists is limited by their prior investments in knowledge domains, which in scientifically disparate teams reflects itself as a problem in the assimilation and implementation of external knowledge (Zahra *et al.*, 2002). Consequently, I expect a curvilinear relationship between scientific proximity in an inventing team and the impact of inventions produced, such that moderate distance is better than low or high distance for the impact of inventions created in a technological field.

> *Hypothesis 2: The relationship between the scientific proximity across inventors of a focal patent and the impact of the focal patent is curvilinear (inverted U-shaped) such that moderate scientific proximity creates the highest impact.*

2.2.3 Coordination across Departments

Coordination costs emerge when individuals or organizations engage in collaborations from anticipated organizational complexity of decomposing tasks among partners, the ongoing coordination of activities to be completed jointly or individually across organizational boundaries, as well as pressing communication needs (Gulati *et al.*, 1998). Concerns about coordination costs are significant in settings which can involve significant coordination of activities between the involved parties, but have to be managed without the benefit of the structure and systems available in traditional hierarchies (Litwak *et al.*, 1962). Due to the specific nature of research in stem cells which includes collaborative work of scientists who participate from different geographic, institutional and departmental settings and whose activities are not governed by a hierarchical control system, stem cell research constitutes a context in which the concerns about coordination costs are valid.

In settings which involve joint project work, organizational theorists have defined four coordination activities as being important to integrate and utilize the knowledge of the team to

the best level possible (Cummings *et al.*, 2007). Assigning specialists to appropriate tasks is one of the important coordination activities that reduce over-dependency and communication failures (Weick, 1997). Reducing efforts allocated to communication and information transfer through sharing resources is another coordination activity of importance, while learning and transferring knowledge across team members also brings about the synergistic benefits of knowledge sharing (Cummings *et al.*, 2007). At the same time, as it helps in building trust, enhancing participation and developing respect among team members, direct communication is seen as another critical coordination activity (Cummings *et al.*, 2005, 2007, 2008).

Such activities are important for integrating and utilizing knowledge across the members of a team; however they are not easy to achieve especially in joint project contexts where the involvement of individuals from multiple departments increases the complexity and difficulty of coordination (Cummings *et al.*, 2007; Hobday, 2000). Several factors contribute to increasing the coordination costs in research projects that include multiple departments. Involvement of researchers across multiple departments decreases the chances of building common ground (Clark *et al.*, 1991), maintaining awareness of what others are doing (Weisband, 2002) and being able to adjust to surprises (Cummings *et al.*, 2007).

Hypothesis 3: The effect of the number of unique departments represented on a patent team on the scientific impact of the focal patent is negative.

2.2.4 Scientific Proximity and Generalized Experience

Of the two essential parts to the knowledge recombination problem taking place in a creative team (Taylor *et al.*, 2006); scientific proximity determines the recombination space over which scientists can engage in diverse approaches to solving problems, drawing on sets of knowledge that are distant from one another, while prior experience helps develop the common knowledge across inventors and improves task allocation. However, when teams have prior

experience from a context outside of their current domain, this experience is reflected as formation of routines and processes that may not fit with the necessities of the changing environment.

Search for new knowledge is a path dependent process (Song *et al.*, 2003), shaped by past experiences (Nelson *et al.*, 1982). Success in the past may result in reduced incentives to experiment new ways of doing things (Sørensen *et al.*, 2000) and development of routines becomes more standardized (Song *et al.*, 2003). Existing routines stemming from successful collaborations in the past decrease the ability to adapt to the requirements of change. The existence of outside domain experience interacts with the ability to recombine knowledge such that, for those teams with low scientific distance, (cognitive distance) (Nooteboom *et al.*, 2007), the impact of their work is decreasing as the scientific distance increases because the marginal effect of communication problems is higher than the marginal effect of recombination benefits which are reduced due to the existence of routines from earlier regime. At the same time for those teams at medium levels of scientific distance, the impact of their work increases as their scientific distance increases because the effect of recombination benefits are more pronounced than the effects of common language problems since the language problems are resolved in previous work experience. Hence prior recombination ability of a team may work against the team when a contextual change occurs resulting in a shift in the fit of joint knowledge of scientists with the environment.

> *Hypothesis 4: For teams with high generalized experience, scientific proximity will have a curvilinear (U-shaped) relationship to impact such that when the proximity is low/high the invention of the team will have a higher impact than when the proximity is at medium levels.*

2.2.5 Coordination across Departments and Generalized Experience

Multi-departmental teams are known to suffer from problems of coordination and effective distribution of their joint work (Cummings *et al.*, 2005, 2007). Teams with multi-departmental structures can better mitigate their coordination problems stemming from the multi-departmental structure of their team when they have prior joint work experience. Such experience may be an essential factor in helping bridge the coordination gap through development of transactive memory that allows team members to effectively communicate and distribute tasks across the team. A team with prior joint work experience may be able to overcome the coordination problems (Bercovitz *et al.*, 2011; Cummings *et al.*, 2007; Taylor *et al.*, 2006) to reap the benefits of inter-departmental collaboration.

> *Hypothesis 5: In teams with high generalized experience, the relationship between the number of departments represented on a patent team and the scientific impact of the focal patent is positive.*

2.3 Method

2.3.1 Sample and Data

To examine new knowledge creation and the impact of such knowledge on future inventions, I use patents as indicators of innovation efforts. Patents provide an excellent trail of codified knowledge and have been widely used in the context of innovation studies involving knowledge recombination and spillovers (Agarwal *et al.*, 2009; Ganco, 2008; Jaffe *et al.*, 1993). I am interested in analyzing the collaboration benefits and coordination costs of a focal team on the technological impact of the team's work. The unit of analysis is the patent, while the level of analysis is the patent team consisting of individual inventors of a patent. In my analyses, each patent represents a knowledge creation effort by a team of designated individuals.

The sample to test the hypotheses consists of 314 patents granted in human embryonic stem cell research during the period between 1998 and 2010. I retrieved these patents by searching relevant key words (human and embryonic stem cells) in the United States Patent and Trade Office (USPTO) and European Patent Office (EPO). I then downloaded all the available information in the patent. I supplemented this data by retrieving inventor affiliations from the corresponding author institutions that were disclosed in scholarly article publications on or before the date at which the application for the focal patent was filed. I gathered additional information from the inventors from their designated web pages and from their CVs by going through each inventor manually. To further extend the dataset, we matched the patent data with publication data available in ISI database. Different individuals may use different names, so I accounted for different spellings to ensure accuracy. Further, I used information on whether or not the individuals in the teams had gained access to the scarce stem cell lines available at the Wisconsin Alumni Research Foundation. Due to the legal environment at that time, federal funding was available only to those pre-approved lines (Jain *et al.*, 2007); therefore having access to and experience working with stem cell lines could prove to be a valuable resource.

2.3.2 Measures

Dependent Variable.

Scientific Impact. To assess the impact of the patent and the underlying invention, I used the cumulative forward citations to an individual patent. Forward citations count the number of times a patent (the "cited patent") is included in the prior art of subsequent patents. I recorded the total number of forward citations a patent received from the time it is granted until the end of the study period as an indicator of the impact of invention. These citations come from the entire

population of patents in USPTO and EPO, which also includes the sample of 314 patents used in this paper.

Prior work has noted that the effects stemming from having diverse technological fields such that patents in crowded fields may be cited more than patents in sparse fields simply because the population of citing patents is higher (Gittelman *et al.*, 2003). The data set is already limited by its focus on hESC patents and thus largely accounts for the effects of technological crowding. Each patenting activity by a team shows evidence of a collaborative activity. It is the responsibility of the inventor to cite appropriate prior art (patents granted earlier that are relevant to the invention). Such citations need the approval of the patent examiner and this approval helps in removing inventor bias to a considerable extent from citation behavior. As previous research suggests, forward citations to a patent can be considered a measure of technological impact (Albert *et al.*, 1991) and a proxy for economic value to the innovator (Hall *et al.*, 2005). Earlier research has suggested that in the life sciences, patents are a crucial means of appropriating returns to innovation and hence in this field citation rates are more likely than in other fields to contain information about the technological and economic value of a given invention (Gittelman *et al.*, 2003; Powell *et al.*, 1996). I calculated two measures, forward citations and non-self forward citations, which are both closely related. I report the results for forward citations where self-cites are excluded, but the results are consistent when self-cites included as well.

Independent variables

Generalized experience (outside the hESC domain). I calculated the cumulative number of times that a given team of individuals on a patent have worked together previously up to the time the focal patent was filed (Nerkar, 2003; Reagans *et al.*, 2005; Rowley *et al.*, 2000). Teams

collaborating on multiple prior patents are likely to develop a greater understanding of the research and institute better work routines to coordinate their work.

Individual experience. Teams consist of lead inventors as well as other members of a patent team that have or have not previously worked on patents. I refer to the prior patenting experience of the lead inventor (the first inventor listed on a patent) and measure the experience of the lead inventor by counting the number of times s/he has patented in hESC prior to the focal patent.

Number of departments. Using information from the content coding of all inventors' affiliations, this variable refers to the total number of unique departments that are represented on a team. Each individual's departmental affiliation is compared with those of the other members on the team and the total number of unique departments is used as a measure of the dispersion resulting from having multiple departments being represented on the team. For instance, for a team involving five researchers from three different departments, we would give a score of three (Cummings *et al.*, 2007).

Scientific proximity. I measure scientific distance among the patent inventors as calculated from the MeSH classifications of their previous publications. Using "Publication Harvester" program (Azoulay *et al.*, 2006) and the "Scientific Distance Report" application embedded, I gathered all publications for each inventor and calculated the propensity of overlap among inventors using MeSH terms assigned to each publication of each inventor in the PubMED database. For any of the two inventors in the team, this measure uses published papers to calculate the following: the number of total MeSH classifications that are used to classify both the focal inventor and the tie at year t-1, divided by the total number of MeSH classifications of focal inventor's publications at year t-1, t being the year of patent application for the focal patent.

This is a continuous measure between 0 and 1, as the scientific proximity among patent inventors increases (i.e. the overlap across their MeSH classification increases), this value gets closer to 1. I use the distance between lead inventor and last inventor as representative of the team overall distance since these two individuals have the most important roles in inventive teams.

Control variables

Time elapsed from grant date. The baseline of cumulative forward citations to a patent is influenced by its age, and it is therefore necessary to control for the time elapsed between the granting of the patent and the date that the citation data was collected (Gittelman *et al.*, 2003; Podolny *et al.*, 1995). I develop a variable that counts the number of days between the granting date of the patent and the date that the citation data was collected.

Team size. Team size is likely to influence both the benefits of collaboration as well as coordination costs. Larger teams are likely to expend more effort in coordinating actions whereas these larger teams also increase access to a broader set of skills. Team size acts as a proxy for the resources invested in the research project which may affect the research outcome (Gittelman *et al.*, 2003; Podolny *et al.*, 1995). Teams account for a larger proportion of science and their work achieves a higher impact than lone inventors (Wuchty *et al.*, 2007). Team size is measured as the total number of inventors that are working on a patent team.

Number of institutions. I created a variable that counts the number of unique institutions represented on a team of inventors. Teams that involve individuals from multiple institutions face greater obstacles to coordinate their activities.

Joint publications experience. Teams with higher scientific capability, may be better equipped to develop high impact innovations. This may be particularly pertinent in science-based industries such as biotechnology (Powell *et al.*, 1996). To control for this, I matched all

individuals represented on the patents with information about all publications on hESC listed in the Science Citation Index. I went through all names manually to account for different ways of spelling names and to make sure we retrieved the full publication records for the individuals represented on the patents. The team's scientific publications is measured by aggregating the number of publications on which inventors of a patent team are listed as co-authors.

Patent scope. USPTO and EPO use a classification system where each patent is assigned to relevant classes. The number of patent classes that a particular patent is assigned to is seen as a proxy for the breadth of the patent that influences the patent's subsequent impact (Lerner, 1995).

Number of Claims. Claims in a patent are argued to provide information about the intellectual space that the patent protects (Lanjouw *et al.*, 2001). Thus, patents with more claims may have a higher likelihood of getting future citations. To control for this, I include a measure of the number of unique claims made in each patent.

Geographic Distance. I have calculated the geographic distance among the members of a scientific team using their affiliation information and the coordinates of their labs. The geographic distance measure calculates the distance across inventors on Euclidean space with the final distance measure calculated as an average of all distance across inventors.

Multiple Affiliations Dummy. I created a dummy variable taking the value of 1 if the patent team involves any inventor(s) who are affiliated with more than one institution in the year before filing the invention.

2.3.3 Analysis

The dependent variable is the count of non-self forward patent citations, which is heavily skewed with many observations receiving a low number of citations. Such count data are usually

estimated with one parameter Poisson models, but because of over-dispersion, Poisson estimates may be biased (Cameron *et al.*, 1986). I therefore employ negative binomial regression models to correct for this potential bias. In my analysis, I ran alternative specifications using zero-inflated negative binomial as well as Poisson models and the results were stable across models. I report the negative binomial regression results in the following section.

2.4 Results

In Table 1, I report the descriptive statistics and correlations of the variables. Although the correlations reported are relatively low, I derived the variance inflation factors (VIFs) as a control from the final model and the average VIFs were well below the general accepted threshold of 10 (Greene, 1997). VIFs are typically made as a post estimation command for linear regressions, but as they make an assumption about the relationship between the independent variables, it is possible to use this technique for other functional forms as well (Menard, 2002). In testing the hypotheses, I also included the variables stepwise to check and make sure that the signs of coefficients are stable across the regressions. If multicollinearity would have been a major issue, sign and coefficients could have changed direction. In testing the interaction effects, I also mean centered the variables (Aiken *et al.*, 1991). Taken together, these precautions minimized the problem of multicollinearity.

TABLE 1: DESCRIPTIVE STATISTICS AND CORRELATIONS

	Variables	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	Non Self cites	3.61	16.51	0.00	260.00	1.00																	
2	Time since granting	1785.85	851.64	110.0	4529.0	0.30	1.00																
3	Number of inventors	3.33	2.04	1.00	11.00	-.06	-.11	1.00															
4	Number of institutions	1.70	1.01	1.00	6.00	-0.02	0	0.52	1.00														
5	Number of patent classes	6.21	5.19	0.00	24.00	0.23	0.55	-0.03	0.06	1.00													
6	Multiple Affiliations Dummy	0.26	0.44	0.00	1.00	0.03	0.01	0.01	-0.01	-0.01	1.00												
7	Joint Publication Experience	0.30	0.46	0.00	1.00	-0.04	-0.05	-0.34	-0.28	-0.07	0.03	1.00											
8	Claims	34.84	38.13	1.00	372.00	0.02	-0.01	-0.01	-0.04	0.01	0.09	-.05	1.00										
9	Generalized Experience	0.29	0.46	0.00	1.00	-0.02	0.05	-0.32	-0.11	0.01	-.08	0.34	0.02	1.00									
10	Geographic Distance	780.57	2558.37	0.00	16929	-0.01	0.23	0.02	0.22	0.19	-.03	-.10	0.03	-0.10	1.00								
11	Scientific Proximity	0.36	0.30	0.00	1.00	0.03	0.11	-0.36	-0.31	-0.01	-.01	.50	-0.01	0.39	-0.10	1.00							
12	Proximity Squared	0.22	0.31	0.00	1.00	0.01	0.10	-0.39	-0.31	-0.02	0.01	0.52	-0.03	0.44	-0.10	0.96	1.00						
13	Number of Departments	1.92	1.20	1.00	10.00	-0.04	-0.07	0.56	0.87	0.02	0.10	-.24	-0.06	-0.17	0.15	-0.32	-0.32	1.00					
14	Institutional Multiplicity	0.15	0.35	0.00	1.00	-0.01	-0.03	0.18	0.49	0.02	0.02	-.17	0.08	-0.03	0.13	-0.15	-0.15	0.37	1.00				
15	Lead inventor Experience	2.84	3.90	1.00	28.00	-0.04	-0.13	-0.13	-0.13	-0.11	0.23	0.07	-0.12	0.05	-0.03	-0.05	-0.01	-0.11	0.37	1.00			
16	Sci. Proximity X Generalized Exp	0.16	0.32	0.00	1.00	0.00	0.12	-0.42	-0.24	0.03	-.05	0.52	-0.04	0.77	-0.11	0.74	0.81	-0.28	-0.11	-0.09	1.00		
17	Sci Prox Square X Generalized Exp	0.13	0.31	0.00	1.00	0.01	0.12	-0.41	-0.25	0.02	-.02	0.52	-0.06	0.64	-0.11	0.77	0.86	-0.28	-0.28	-0.12	0.08	1.00	
18	Number of Dept X Generalized Exp	0.47	0.93	0.00	6.00	0.01	0.01	-0.10	0.22	-0.01	-.05	0.11	0.01	0.78	-0.04	0.10	0.13	0.17	-0.28	-0.14	0.06	0.97	1.00

TABLE 2: NEGATIVE BINOMIAL MODELS OF INVENTION IMPACT

Variable	Model 1	Model 2	Model 3	Model 4	Model 5
Time elapsed	0.00163***	0.00173***	0.00167***	0.00175***	0.00167***
	[0.000160]	[0.000154]	[0.000151]	[0.000155]	[0.000154]
Number of inventors	-0.0124	-0.0406	-0.00747	-0.0476	-0.0148
	[0.0630]	[0.0647]	[0.0645]	[0.0650]	[0.0648]
Number of Institutions	-0.251**	0.0583	-0.124	0.0273	-0.137
	[0.109]	[0.237]	[0.245]	[0.234]	[0.242]
Number of Patent Classes	0.0392*	0.0422**	0.0486***	0.0427**	0.0491***
	[0.0211]	[0.0193]	[0.0188]	[0.0187]	[0.0185]
Multiple Affiliations Dummy	0.0476	0.0438	0.0266	-0.0633	-0.0625
	[0.225]	[0.236]	[0.228]	[0.239]	[0.231]
Joint Publications Experience	-0.367	-0.146	-0.105	-0.153	-0.151
	[0.228]	[0.250]	[0.246]	[0.257]	[0.251]
Number of Claims	0.00842***	0.00968***	0.00904***	0.0101***	0.00929***
	[0.00287]	[0.00295]	[0.00287]	[0.00301]	[0.00292]
Generalized Experience		-0.629**	-1.461***	0.792	-0.189
		[0.261]	[0.425]	[0.647]	[0.768]
Geographic Distance		-6.06E-05	-4.34E-05	-6.29E-05	-4.55E-05
		[4.04e-05]	[4.03e-05]	[3.93e-05]	[3.95e-05]
Institutional Multiplicity		0.851**	0.881***	0.918***	0.935***
		[0.338]	[0.328]	[0.337]	[0.326]
Lead Inventor Experience		0.0812***	0.0772***	0.103***	0.0972***
		[0.0272]	[0.0267]	[0.0285]	[0.0280]
Number of Departments		-0.365*	-0.450**	-0.312	-0.395*
		[0.215]	[0.215]	[0.212]	[0.213]
Scientific Proximity		3.084**	2.363*	5.290***	4.933***
		[1.334]	[1.331]	[1.906]	[1.877]
Scientific Proximity Squared		-2.880**	-2.015	-4.996**	-4.781**
		[1.274]	[1.284]	[2.137]	[2.099]
Sci. Proximity X Exp outside hESC				-7.875**	-7.145**
				[3.340]	[3.249]
Sci. Proximity Squared X Exp outside hESC				6.786**	6.650**
				[3.189]	[3.095]
Number of Departments X Exp outside hESC			0.461**		0.426**
			[0.194]		[0.195]
Log Likelihood	-517.8	-501.5	-498.5	-498.6	-496.1

* p<.1, ** p<.05 ***p<.01; two-tailed tests. Robust standard errors are in brackets. 314 hESC patent observations.

TABLE 3: VARIABLE DEFINITIONS

Variable Name	Variable Definition
Control Variables	
Time Elapsed	The number of days that have passed since the inventors have filed their patent application.
Number of Inventors	The number of inventors listed on the patent as the patents authors.
Number of Institutions	The number of unique institutions represented on the patent as calculated from the institutional affiliations of patent inventors.
Number of Patent Classes	The number of unique designated international patent classes which the patent belongs to.
Multiple Affiliations	A dummy variable taking the value of 1 if the patent team has a member who has affiliations in multiple departments at the time of patent filing.
Joint Publication Experience	A dummy variable taking the value of 1 if the patent inventors have previous joint work experience as can be tracked from their joint work in publications before the focal patent is filed.
Number of Claims	The number of claims to novelty made by the patent in the patent document.
Independent Variables	
Experience working outside hESC	A dummy variable that takes the value of 1 if the patent inventors have previous joint work experience as can be tracked from their joint work in patents outside the hESC realm before the focal patent is filed.
Geographical Distance	The geographical distance in miles between the members of the patent team as calculated from the locations of their institutional affiliations at the time of the filing of focal patent.
Scientific Proximity	A measure of scientific distance among the patent inventors as calculated from the MeSH classifications of their previous publications. It's a continuous measure between 0 and 1, as the scientific proximity among patent inventors increases (i.e. the overlap across their MeSH classification increases), this value gets closer to 1.
Scientific Proximity Squared	The squared term for scientific proximity measure.
Number of Departments	The number of unique departments represented on a patent team as calculated from the institutional affiliations of a patent's inventors.
Institutional Multiplicity	A dummy variable taking the value of 1 if a patent team has inventors who are affiliated with industry as well as inventors who are affiliated with public research organizations including at the same time. If all inventors are from industry or all inventors are from public sector this variable takes a value of 0.
Lead Inventor Experience	The experience of lead inventor in patenting in hESC from previous patents filed in the hESC area.

Table 2 shows the results of the negative binomial regression analysis where the dependent variable is the counts of patent citations. Model 1 shows the baseline model which includes controls for time elapsed since the patent was granted, team size, number of claims, patent scope, the team's scientific publication record and access to stem cell lines variable. In Model 2, I introduce three of my theory variables of interest (experience outside the hESC, scientific proximity and number of departments). Model 3 adds the interaction terms between the number of departments and experience while Model 4 includes the interaction terms between scientific proximity and experience while Model 5 includes all variables of concern. Each model represents a significant improvement over the baseline model with the log likelihood value increasing from –517.8 for the base model ($p<.001$) to –496.1 ($p<.001$) for the final model represented in Model 5 of Table 2. The baseline model was generally consistent with prior research findings for patent variables. In line with previous research findings, the number of days in between variable which accounts for the age of a patent has a positive effect on the citations received by the patent (Gittelman *et al.*, 2003).

Hypothesis 1: *Generalized experience outside hESC.* I predicted that the generalized experience, which is defined as having joint work across team members in domains outside of hESC, will have a negative effect on the impact of the invention. In Model 2, the coefficient for experience outside variable was statistically significant. The main effect of the experience outside variable was negative ($b=-0.629$, $p<.05$), this effect remains significant in the Model 3 as well ($b=-1.461$, $p<.01$). This result lends support to the hypothesis that prior experience is detrimental when environment changes.

Hypothesis 2: *Scientific Proximity.* I argued that scientific proximity will have a curvilinear (inverted-U shaped) effect on the impact of inventions generated. In Model 2, I include the scientific proximity (b=3.084, p<.05) and scientific proximity squared (b=-2.880, p<.05) terms and observe that there's a diminishing effect of scientific proximity on impact. Hypothesis 2 is supported.

Hypothesis 3: *Coordination across Departments.* I predicted that the number of departments that are represented on a patent team could have a negative effect on the impact of the invention. Given Model 2's improvement over prior model, I inspect the coefficient of number of departments variable which is significant (the coefficient is significant; b=-0.365, p<.1). The result holds consistently in Model 3 (b=-0.45, p<.05), this variable does indeed point to the negative effect of number of departments represented on a patent team, thereby supporting Hypothesis 3.

Hypothesis 4: *Interaction between generalized experience and scientific proximity.* I argued that in the existence of prior experience, scientific proximity has a larger effect on the impact of inventions. In Model 4, I include the interaction effects between prior experience and scientific proximity and the interaction between prior experience and scientific proximity squared variables. Unlike my expectations though, the interaction effects do not shift the proximity to impact curve upwards but cause a change in shape of the curve, transforming the curve into a U-shaped structure. Both interaction effects were significant, (b=-7.87, p<.05, b=6.78, p<.05). In order to interpret this interaction effect, I followed the suggestions by Aiken and West (1991) and visually graphed the effect. Figure 1 illustrates the effect; in the existence of prior experience the impact of scientific experience on patents exhibits a U-shaped curve where medium levels of proximity have the highest effect on the impact of patents.

Hypothesis 5: *Interaction between prior experience and number of departments.* I argued that the benefits of joint production experience will be to reduce the potential coordination problems that teams having inventors from multiple departments may face in terms of communication and task allocation. In Model 3, I include the interaction effects between joint production experience and the number of departments. The interaction effect was positive and significant as expected, lending support to my hypothesis ($b=0.461$, $p<.05$). In order to interpret this interaction effect, I followed the suggestions by Aiken and West (1991) and visually graphed the effect. Figure 2 illustrates the effect.

FIGURE 1- The Effect of Scientific Proximity and Outside Experience on Impact

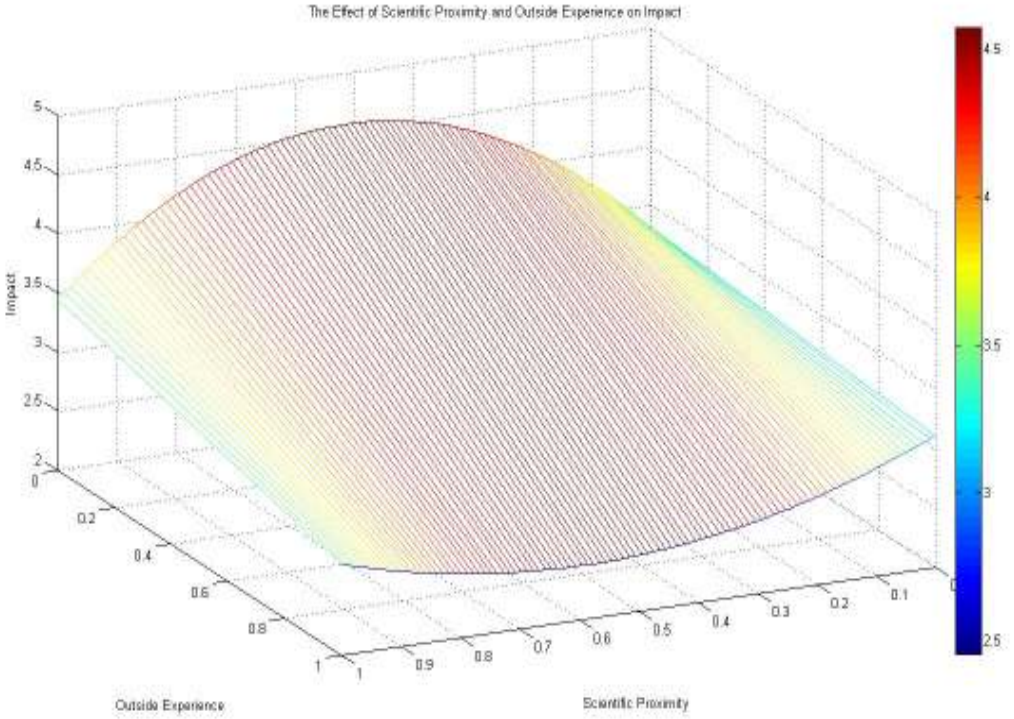
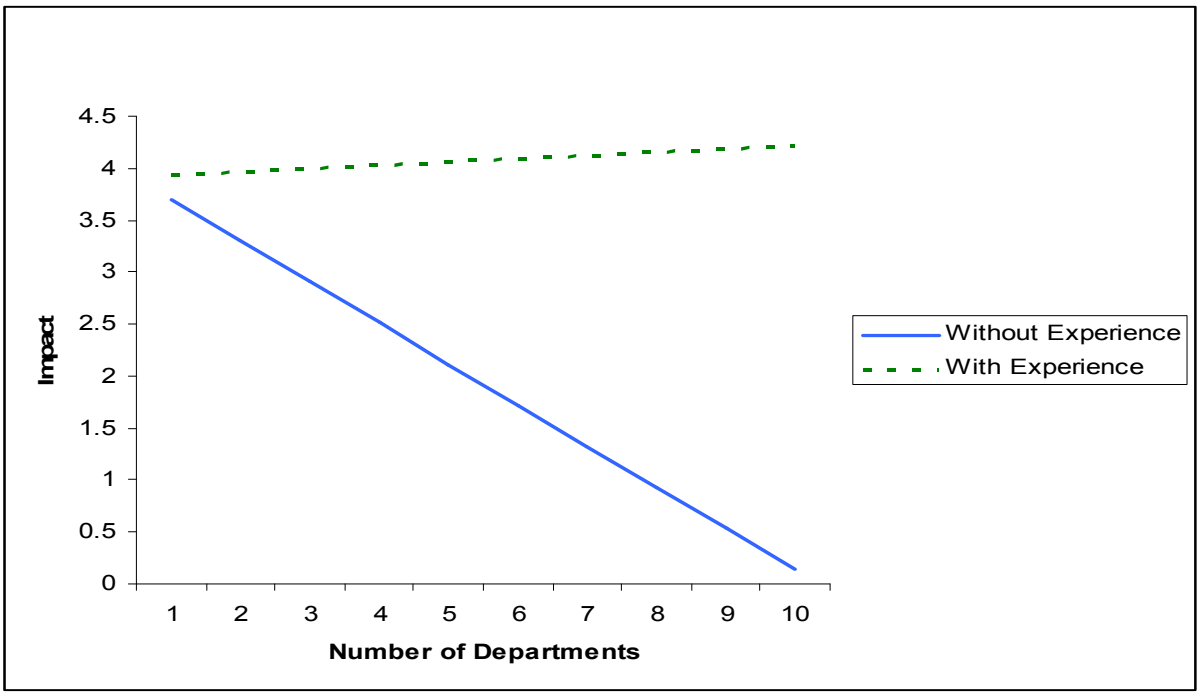


FIGURE 2-The Effect of Number of Departments and Outside Experience on Impact



2.5 Discussion

The role of teams in the discovery process is understudied, even though recent studies systematically document that collaborative teamwork in science has increased substantially over the past few decades (Guimera *et al.*, 2005; Jones *et al.*, 2008). The primary issue is one of harnessing the benefits of collaboration (by knowledge recombination and sharing experience) while reducing the costs of coordination (by reducing costs associated with managing interdisciplinary and geographically dispersed teams). If one could enhance collaborative benefits and reduce costs simultaneously, the value created from collaboration through discovery would likely be enhanced greatly. This study documents how the benefits of collaborative experience make a difference on the scientific impact of the discovery itself. Further, I find evidence that prior experience has diverse effects on mitigating coordination costs stemming from scientific distance (scientific distance across individual inventors on a patent team) and departmental distance. Below, I explain the implications of these findings.

First, I supply empirical evidence for the effects of joint collaborative experience and its attendant coordination problems on scientific impact in a radical science such as stem cells. My primary contribution lies within the emphasis of coordination costs that constitute an important barrier on the effectiveness and efficiency of joint work in innovative settings. A plethora of research has documented the benefits that accrue from collaborating with partners (Cummings *et al.*, 2005, 2007; Fleming *et al.*, 2007; Guimera *et al.*, 2005; Jones *et al.*, 2008; Powell *et al.*, 1999; Powell *et al.*, 1996). A much less attended issue that I attempt to bring to the fore is the coordination costs that emerge when there are geographical, institutional and cultural barriers to bridge that require ongoing dialogue and negotiation (Grant, 1996). By combining the

collaborative benefits and coordination costs, one can reach a more fine-grained understanding of under what conditions collaborations may succeed.

Second, I was able to analyze the effects of coordination costs on the research outcomes of collaborative work in human embryonic stem cells. Having inventors from different departments on a patent team increases the costs of coordination as it increases the complexity and difficulty of communication, task allocation and effective management of a joint work project. It is harder for teams with inventors from multiple departments to bridge institutional gaps and have a shared social setting or maintain awareness of what others are doing. However, in the existence of prior work relationships, coordination problems are overcome and departmental diversity contributes to knowledge creation activity of the team. My findings build up on previous literature and further tease out the effects of coordination problems that hamper interdepartmental collaborations (Cummings *et al.*, 2007). Attending to such problems is important as the number of interdepartmental teams has increased steadily over the last 25 years (Jones *et al.*, 2008). This trend is driven by several different motives including access to new instrumentation (de Solla Price, 1986), reduced communication costs enabled through new communication technologies (Agrawal *et al.*, 2008), complementarities of knowledge and experience to generate new scientific insights (Basalla, 1988) and solving increasingly complex problems that would be intractable for a single individual to solve. Despite these important observations, scholars have paid significantly less attention to the potential coordination costs that emerge from such interdepartmental teams and its affect on scientific impact. Given the results of this study, I believe this is an unattended issue of importance for understanding the conditions under which interdepartmental teams will be successful.

Third, heterogeneity among the inventors on a patent team was expected to have a positive and diminishing effect on the impact of the resulting invention because the ability of the team to recombine knowledge peaks when the there is enough absorptive capacity across the team so that members can understand each other's work while being distant enough to bring in new perspectives to the solution domain. (Owen-Smith *et al.*, 2003). The results show that the scientific distance among assignees of a patent does have an inverted-U type curvilinear shape as expected. However, the interaction effects point to differential effects of scientific distance in the existence and in the lack of prior experience outside the hESC domain. I observe that when there is no prior experience, the scientific distance has an inverted U shaped effect on the impact of the invention being generated. However, when team members have prior joint work experience outside of hESC domain, the effect of distance is a U-shaped curve where either low or high levels of distance are optimal. This may be because in the existence of prior experience outside hESC, teams build work routines that aren't applicable to hESC domain; the existence of such routines creates rigidities which moderate the relationship between scientific distance and knowledge creation in hESC. This is important for purposes of policy makers, as well as researchers and private funders of research since the value of a patented invention is signaled by citation impact and this impact is driven by the structure of scientific team undertaking the research.

Finally, I find evidence for positive effects of prior experiences in patenting in a given area, hESC in this case, are hampered when the activities of the research team have to be coordinated across multiple departments and across geographic space. The learning literature has made great efforts to investigate how individuals and organizations draw inferences from past experiences (Levitt *et al.*, 1988). My findings suggest that experiential learning is greatest when

the team is co-located. Compared to interdepartmental teams, local teams can arguably spend more time discussing and resolve potential conflicts that arise. By working in the same department, tacit knowledge may spill over to other team members enhancing the possibility for the team to improve.

2.5.1 Limitations

In spite of this study's contributions, limitations exist. I followed the tradition of learning curve research of studying the actual outcomes of learning and prior experience rather than measuring directly intermediate processes and mechanisms (Arrow, 1962; Yelle, 1979). In my study, I measured experience outcomes related to learning on the impact of patents produced, but due to data limitations, I could not have intermediate level data to actually measure the development of collaboration routines that would be taking place during joint work at the team level. While I do control for joint work outside and inside the hESC domain in patenting and in publishing, I do not have a control for joint work in labs that's not documented in an outlet. Future research could make use of self reported research endeavors to mitigate the concerns related to the structure of teams whose joint work aren't documented in patents or publications. Also, data on licensing and revenue generation are not available in my context primarily because of the early stage of developments involving therapeutic applications using stem cells; hence my measure of patent impact remains the only major determinant of invention success. At the same time, an important attribute of this context is the existence of and dependence on cell lines, the use of which requires expertise particular to these lines. In settings where such platforms are not necessary or where the knowledge of their operation is orthogonal to changes in the system, it would be possible to observe dissimilar effects of prior expertise on the success of focal work.

2.6 Conclusion

Taken as a whole, the results proffer evidence that team level joint production experience and team structure in terms of scientific and departmental distance are important determinants of patent impact. Having scientists from multiple departments reduces the impact of the patented invention; however, having medium level of scientific distance across inventors enhances the positive effect of recombination on patent impact. At the same time, learning from experience appears to be an important factor determining the patent outcomes. More importantly, although much enthusiasm surrounds interdepartmental collaborations, an issue that has been somewhat neglected was the coordination costs that constitute the downside of collaboration. This study attempts to offer a test of when the benefits from collaborations may outweigh the costs of coordination in research leading to patented inventions. Further research should focus on the dynamic processes that are involved in the initiation of joint work and how collaborations develop despite coordination problems. Another thread of work is needed to uncover the underlying mechanisms by which the different ownership structures governing the patents influence the coordination of activities and affect the impact of the resulting patents.

CHAPTER 3

Complexity and Coordination in Distributed Work

3.1 Introduction

Inter-organizational coordination of activities has been the focus of management scholars for a substantial amount of time. Within this stream of research, scholars have analyzed the role of trust, organizational structure, learning and communication as important determinants of effective coordination mechanisms in inter-organizational collaborations (Ahuja, 2000; Gulati *et al.*, 1998; Gulati *et al.*, 1999; Krishnan *et al.*, 2006; Zollo *et al.*, 2002).

Within this literature particular attention has been paid to the coordination of inter-organizational work in distributed work projects (Cummings *et al.*, 2005; Espinosa *et al.*, 2007; Hinds *et al.*, 2005; Kraut *et al.*, 1999; Srikanth, 2010). One of the factors that determine the success of such projects is the design of the project in terms of how tasks are defined, how procedures are set and how project distribution is allocated across locations.

Coordination through planning mainly deals with task decomposition and modularization in the design stage, in order to avoid the necessity of on-going communication for coordination of action (Galbraith, 1977; March *et al.*, 1958; Tushman *et al.*, 1978). This approach has been analyzed mainly in large scale projects from a modularization and task decomposition perspective.

Although, researchers have touched upon the importance of coordination mechanisms and the effective management of inter-organizational joint work in partnerships (Powell *et al.*, 1996; Rothaermel *et al.*, 2006; Rothaermel *et al.*, 2007), they have mainly focused on the choice and selection of partners (Dyer *et al.*, 1998; Gulati *et al.*, 1998) and the ability to manage joint

work through learning and capability development (Hoang *et al.*, 2005; Kale *et al.*, 2002; Kale *et al.*, 2007).

However, an important set of coordination activities that has been pretty much ignored takes place during the initial stage of distributed work projects where organizations design the project and set the task characteristics and the associated work procedures (Azoulay, 2004; Huckman *et al.*, 2008). Design and planning stages are essential in coordinating distributed work projects (Galbraith, 1977; Kiesler *et al.*, 2002; March *et al.*, 1958; Tushman *et al.*, 1978), yet the design complexities that affect the effective coordination of distributed work projects has been little studied. Hence, in this study, I'll be exploring the extent to which complexities in the design of distributed work projects stemming from task size, task breadth and partner complexities affect the coordination of distributed work and the performance outcomes.

3.2 Theory and Hypotheses

3.2.1 Task complexity and Coordination

One of the factors that can make a project complex is the sheer size of the task involved (Espinosa *et al.*, 2007). Although size in terms of the number of instructions has been shown to be an important factor determining the complexity of projects in settings such as software development (Banker *et al.*, 1998; Espinosa *et al.*, 2007) and patent examination (Harhoff *et al.*, 2009), task complexity can have many other attributes in different settings.

One such attribute is the procedural complexity of clinical trials that stems from the inclusion of a number of interventions that the trial involves. Interventions in a clinical trial include use of drugs, genes, vaccines and devices as ways of changing the status of subjects' health to bring out a desired outcome (FDA, 2010). As the number of interventions in a clinical study increases, the procedural complexity of carrying out the trial surges. One reason for this is

the above mentioned complexity associated with the size of the task; more instructions listed in the protocol require a greater knowledge and understanding of how to proceed with the interventions and what to do if unexpected outcomes arise.

At the same time, the temporal dimension of procedural complexity comes to the fore as multiple interventions are administered at the same time. The ability to oversee multiple interventions and to disentangle the individual effects of these interventions becomes excessively hard as the number of interventions increases. Moreover, the substance of interventions vary across treatments and there may be complementarities across interventions which require special attention, i.e. a drug that regulates insulin which should be applied while a high fiber, low calorie diet should be maintained. In other words task complexity may not only increase with task size but also with the amount of information and number and types of interrelationships between sub-processes related to these tasks (Espinosa *et al.*, 2007). This is similar to the concept of coordinative complexity discussed in literature (Wood, 1986). The increasing level of task complexity stemming from the involvement of multiple procedures generates coordination problems and negatively effect project performance in distributed work.

Hence, I hypothesize that;

Hypothesis 1: The effect of the number of interventions that are involved in a clinical trial on the timely completion of the clinical trial is negative.

3.2.2 Task breadth and variability

Task breadth is another factor influencing the effective management of distributed work projects. The literature suggests that as the number of desired outcomes of a project increases, the associated complexity increases to the extent that the outcomes are not positively related with one another (Campbell, 1988). In the case of clinical trials, number of outcomes desired is

associated with the number of conditions that are being addressed in the study. The conditions addressed in a clinical study can be defined as the diseases and disorders for which the given treatment(s) are being tested.

Increase in the breadth of project tasks not only increases the number of desired outcomes, i.e. task dimensions that require attention, but also leads to higher complexity in understanding outcomes of the project because of the potential confounding effects of interactions across outcomes. There is a positive relationship between having an organizational focus on specific tasks and positive outcome in terms of higher output and productivity (Huckman *et al.*, 2008).

For the above mentioned reasons, I expect a negative relationship to exist between task breadth and project performance.

Hence, I propose that;

Hypothesis 2: The effect of the number of conditions addressed in a clinical trial on the timely completion of the clinical trial is negative.

3.2.3 Task applicability

Task applicability refers to the extent to which tasks can be applicable to a greater set of subjects. Projects that involve tasks which can be applied to a greater variety of subjects are projects that show a greater degree of routinization. Such projects need less manipulation and can be more efficiently implemented. The efficiency that arises from task applicability manifests itself positively on the speed of project implementation.

In the clinical trial setting, one measure of applicability is the extent to which studies can involve patients from diverse age groups (FDA, 2010). Studies that allow for a wider population range to be targeted by the treatment involve a higher degree of available routines for conducting the tasks as the tasks are applicable to different subject characteristics in relation to the subject's

age. This increases the level of efficiency in implementing the tasks and results in higher project performance.

Therefore, I hypothesize that;

> *Hypothesis 3: The effect of task applicability as shown by the age range of population that can be included in a clinical trial on the timely completion of the clinical trial is positive.*

3.2.4 Partner Complexity and Coordination

Coordination costs include anticipated organizational complexity of decomposing tasks among partners, the ongoing coordination of activities to be completed jointly or individually across organizational boundaries as well as the communication and decisions that would be necessary to carry out such activities (Gulati *et al.*, 1998). In settings where complex and overlapping labor requires continuous adjustments between partners and forces them to link specific activities with each other closely and regularly, a high degree of interdependence exists (Gulati *et al.*, 1998).

Clinical trials are sponsored by organizations which have to coordinate their responses and make adjustments to the trial process on an ongoing basis as they receive information from the trial sites with regards to the performance of proposed treatments. The trial process may be interrupted by unexpected events such as adverse reactions that require adjustments to the protocol (Azoulay, 2004). As these adjustments require joint decision making given the stakes of sponsors in clinical development, coordination of action becomes an important issue. The involvement of multiple sponsors increases the complexity in managing the trial process and the associated difficulty of coordination by increasing difficulty of joint decision making (Cummings *et al.*, 2007; Hoang *et al.*, 2005) which in turn decreases the reaction time adversely affecting the performance of clinical trial.

Therefore, I hypothesize that;

Hypothesis 4: The effect of the number of sponsors in a clinical trial on the timely completion of the clinical trial is negative.

3.3 Methods

3.3.1 Sample and data

In order to examine the effects of task complexity and coordination costs stemming from the project design in a distributed work environment, I chose clinical trials as a setting. Firms that wish to gain regulatory approval for market introduction of their drugs need to provide substantial evidence of their drugs' effectiveness through controlled clinical trials (FDA, 2010). Clinical trials involve critical tasks such as recruitment of candidates, following up on the protocol set for study, submission of case reports including original patient records and charts (Azoulay, 2004). Clinical trials offer an ideal setting to study design of distributed work projects involving complex tasks because tasks included in clinical trials vary by phase of the drug development process, the conditions addressed, the interventions at use and other trial specifications that determine the routine manipulation, storage and sharing of symbolic information within established categories(Azoulay, 2004; Huckman *et al.*, 2008). At the same time, clinical trials also involve knowledge production which occurs through generation of search rules for identifying problems and heuristics that leads to the resolution problems and which can't be reduced to simple protocol steps (Azoulay, 2004). In this sense, each clinical trial displays a unique set of characteristics which reflects itself on the complexities of tasks involved in clinical research and the interdependencies between the attributes of relevant tasks.

The sample to test my hypotheses consists of 227 diabetes clinical trial observations between the years of 1981 and 2009. I have retrieved clinical trial data from the Federal Drug

Administration's official website that includes data on all clinical trials associated with treatments seeking regulatory approval in the United States. Since United States is the main market for pharmaceutical companies, the FDA database which is the most comprehensive of its sort covers a significant portion of the universe of clinical studies carried across the globe. I have retrieved clinical study data by using relevant search words for identifying studies that are addressing diabetes condition and downloading all available information. After downloading the data, I have gone through 4569 clinical locations associated with the 227 trials listed in the dataset to correct for potential mistypes.

3.3.2 Measures

Dependent Variable.

Timely Completion. To assess the impact of task complexity, breadth and applicability on the timely completion of clinical trials, I have created a timely completion variable consisting of 8 categories in order of their distance in terms standard deviation(s) from the mean completion time of all studies in the dataset. The studies that go into the first category are those which have a completion time that is more than 3 standard deviations above the mean completion time (measured as number of days from the start date of the study to the completion date of the study) of all respective studies in the dataset. The categories follow in order of standard deviation difference, i.e. second category includes those studies that are concluded in a time that is below or equal to 3 standard deviations from the mean completion time, but above 2 standard deviations from the mean completion time. In total this procedure generates 8 ordered categories; the latest completions with respect to the mean constituting the first category and the earliest completion with respect to the mean constituting the last (eighth) category.

In order to overcome the potential problem of mismatches in creating these categories, I have created respective categorical scales for the three phases that are observed in the dataset. For example, Phase I studies are "conducted to determine the metabolism and pharmacologic actions of drugs in humans, the side effects associated with increasing doses, and to gain early evidence of effectiveness" (FDA, 2010) and deal with initial efficacy and safety checks which take significantly less time to complete. Therefore each Phase I study's completion time is compared against the respective mean completion time and standard deviation distances of all Phase I studies rather than the measures for the entire set of clinical trials . I apply the same procedure for constructing the respective categories of Phase II and Phase III studies. Timely completion is an important measure of clinical trial success since sponsors of the clinical trial depend on FDA approval for marketing their drugs (Azoulay, 2004), and potential delays may cause losses in terms of monetary benefits and may even harm the reputation and market position of sponsoring firms.

Independent variables

Task Complexity. I have calculated the number of interventions that are involved in a clinical trial as a measure of task complexity involved in the study. Interventions involve the use of drugs, genes, vaccines and devices in order to bring out a desired change in the health of the patient (FDA, 2010) and as such interventions determine the type and number of tasks to be applied in the clinical trial processes.

Task Breadth. I have calculated the number of conditions that are addressed in a clinical trial study in order to measure the breadth and variability of tasks that are involved in the study. The conditions addressed are diseases and disease related disorders that are being targeted by the interventions in the study. The number of conditions addressed captures the breadth and

variability of tasks involved since the conditions range from cardiovascular diseases to bone diseases, hypercholesterolemia, hypertension, obesity, kidney disorders, diabetic retinopathy and beyond. This variability in the targeted outcome measures of a study reflects itself on the procedures involved such as recruitment, application of treatment and analysis of results in addressing such diverse conditions (FDA, 2010).

Task Applicability. I have calculated the age range determining the recruitment pool desired in a clinical trial study to measure the level of applicability of the tasks on diverse subject pools. Age range is an indicator of the extent to which diverse population groups can be included in the study and how widely the study protocol and the associated treatments can be applied. It measures complexity in terms of applicability and routinization of the tasks involved in the study.

Partner Complexity and Coordination. Using information from the content coding of sponsors collaborating in clinical trials, this variable refers to the total number of unique organizations that support the clinical trial. This variable captures the amount of coordination costs in terms of communication, information sharing and decision making efforts across sponsors of a given clinical trial. It is similar in nature to the measure used to capture cross-institutional coordination in distributed work (Cummings *et al.*, 2005, 2007).

Control variables

Government Organizations. The involvement of government organizations such as "National Eye Institute" or "National Institute for Diabetes and Kidney Disorders" is a factor that may influence the study completion time because these organizations have unique resources such as a grand database of patients registered through government agencies and expertise in conducting clinical trials in house. Hence, to control for the effect of the involvement of such

organizations, I have coded a dummy variable that takes the value of 1 if a government organization is sponsoring the study and 0 otherwise.

Universities and Research Centers. Medical universities and the associated research centers conduct a significant part of their clinical trials in house and their studies are geared toward testing compounds from scientific discoveries for potential applications which creates a bias against clinical studies by these organizations with respect to the studies conducted by pharmaceutical firms. To control for this effect, I introduce a dummy variable taking the value 1 if universities and research organizations are sponsoring the clinical trials and 0 otherwise.

Basic Science Based Studies. Studies designed to test the effect of an intervention based on discoveries of basic science in labs are involved with understanding the underlying mechanisms of the compounds and are not market oriented in nature. Hence, such studies may have a different time frame for application of the study. To control for the effect of such studies, I'm including a dummy variable that takes the value of 1 if the study design involves basic science and 0 otherwise.

Treatment Based Studies. Studies designed to include new treatments which may involve new approaches to surgery or new combinations of drugs rather than existing treatment modes with new compounds are controlled for by a dummy variable that takes the value of 1 if the study is a treatment based study and 0 otherwise.

Randomized Studies. A randomized study is a study in which participants are randomly (i.e., by chance) assigned to one of two or more treatment arms of a clinical trial (FDA, 2010). Occasionally placebos are utilized. To control for potential effects brought about by the modes of conduct between randomized trials and non-randomized trials, a dummy variable is introduced taking the value of 1 if the trial is randomized and 0 otherwise.

Cross-over Assignment Studies. A cross-over assignment refers to the alternating assignment of subjects across placebo and treatment groups over the duration of the study. To control for the effects of implementing this particular design on the timely completion of the given study, I include a dummy variable taking the value of 1 if the study includes a cross-over assignment design and 0 otherwise.

Studies Involving Multiple Phases. A number of studies in the dataset have been designed to address two different phases at the same time. This increases the complexity of the study design. To control for the inclusion of multiple phases in a study, I have created a dummy variable taking the value of 1 if the study involves multiple phases and 0 otherwise.

3.3.3 Analysis

The dependent variable in my analysis is the timely completion of a clinical trial process as determined by the fit of the clinical trial completion time into one of the eight ordered categories that the study duration falls into. As such I am using ordered logistic regression to estimate the effects of my predictors. I also use OLS as a robustness check and control for the differences across phases, I receive similar results. Overall, my results are stable across both approaches with no significant changes in direction, magnitude or significance.

3.4 Results

In Table 4, I report the descriptive statistics and correlations of the variables. Although the correlations reported are relatively low, I have derived the variance inflation factors (VIFs) as a control from the final model and the average VIFs were well below the generally accepted threshold of 10 (Greene, 1997). VIFs are typically made as a post estimation command for linear regressions, but as they make an assumption about the relationship between the independent variables, it is possible to use this technique for other functional forms as well (Menard, 2002).

In testing my hypotheses, I also included the variables stepwise to check and make sure that the signs of coefficients are stable across the regressions. If multicollinearity would have been a major issue, sign and coefficients could have changed direction. Taken together, these precautions minimized the problem of multicollinearity.

Table 5 shows the results of the ordered logistic regression analysis where the dependent variable is the timely completion of clinical trials. Model 1 shows the baseline model which includes controls for involvement of government organizations in the study, inclusion of universities and research centers, study designs involving basic science, new treatments, randomized trials and cross-over assignments as well as a control for studies involving multiple phases. In Model 2, I introduce the first of my theory variables of interest (task complexity). Model 3 includes the second theory variable (task breadth) and Model 4 includes (task applicability). Finally, in Model5, I add the last theory variable of interest (partner complexity). Each model represents a significant improvement over the baseline model and the prior model with the log likelihood value increasing from –308.4 for the base model ($p<.001$) to –289.2 ($p<.001$) for the final model represented in Model 5 of Table 5. The baseline model was generally consistent with prior research findings for clinical trials. In line with previous research findings, the basic science based studies have a negative effect on the timely completion variable (Huckman *et al.*, 2008).

Hypothesis 1: *Task Complexity.* I predicted that the task complexity as represented by the number of interventions involved in a clinical study could have a negative effect on the timely completion of the clinical trial.

TABLE 4: DESCRIPTIVE STATISTICS AND CORRELATIONS

	Variables	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12
1	Timely Completion	4.02	3.14	1	8	1											
2	Government Organizations	.03	.18	0	1	-.06	1										
3	Universities and Research Centres	.30	.46	0	1	-.35	.27	1									
4	Basic Science Based Studies	.01	.10	0	1	-.12	-.02	.07	1								
5	Treatment Based Studies	.85	.36	0	1	.30	-.25	-.31	-.24	1							
6	Randomized Studies	.91	.28	0	1	.04	.07	.07	.04	-.08	1						
7	Cross-over Assignment Studies	.11	.31	0	1	.25	.00	-.02	.21	.09	.12	1					
8	Studies involving Multiple Phases	.08	.28	0	1	.03	.10	.26	.24	-.09	.00	.15	1				
9	Task Complexity	1.62	1.06	1	8	-.11	-.03	-.16	.01	-.04	.19	.06	-.10	1			
10	Task Breadth	1.73	1.37	1	11	-.28	.07	.05	-.01	-.28	.01	-.10	.03	.26	1		
11	Task Applicability	55.03	14.60	11	73	.12	-.17	-.15	.03	.06	-.07	-.09	-.04	.05	.05	1	
12	Partner Complexity	1.55	1.10	1	7	-.38	.31	.49	-.03	-.24	.08	-.09	.14	-.01	.14	-.23	1

TABLE 5: ORDERED LOGISTIC REGRESSION MODELS OF TIMELY COMPLETION

Variable	**Model 1 Timely Completion**	**Model 2 Timely Completion**	**Model 3 Timely Completion**	**Model 4 Timely Completion**	**Model 5 Timely Completion**
Government Organizations	0.592	0.587	0.647	0.857	1.343*
	[0.699]	[0.708]	[0.719]	[0.714]	[0.752]
Universities and Research Centres	-1.395***	-1.583***	-1.627***	-1.604***	-1.080***
	[0.303]	[0.313]	[0.318]	[0.320]	[0.350]
Basic Science Based Studies	-2.367*	-2.482*	-2.797**	-3.010**	-3.360**
	[1.383]	[1.351]	[1.336]	[1.370]	[1.417]
Treatment Based Studies	0.888**	0.816**	0.572	0.566	0.635
	[0.381]	[0.385]	[0.397]	[0.399]	[0.403]
Randomized Studies	0.481	0.766*	0.727*	0.751*	0.791*
	[0.423]	[0.438]	[0.436]	[0.437]	[0.442]
Cross-over Assignment Studies	2.129***	2.338***	2.350***	2.432***	2.366***
	[0.659]	[0.695]	[0.698]	[0.693]	[0.706]
Studies involving Multiple Phases	0.58	0.538	0.605	0.706	0.908*
	[0.495]	[0.503]	[0.500]	[0.504]	[0.525]
Task Complexity		-0.446***	-0.372**	-0.396**	-0.396**
		[0.146]	[0.160]	[0.162]	[0.167]
Task Breadth			-0.338***	-0.355***	-0.362***
			[0.109]	[0.111]	[0.114]
Task Applicability				0.0187**	0.0161*
				[0.00861]	[0.00883]
Partner Complexity					-0.545***
					[0.163]
Log Likelihood	-308.4	-303.5	-298.3	-295.9	-289.2
Wald Chi-Square	59.49***	69.3***	79.67***	84.48***	97.9***

* p<.1, ** p<.05 ***p<.01; two-tailed tests. Standard errors are in brackets. 227 clinical trial observations.

Given Model 2's improvement over prior model, I inspect the coefficient of task complexity variable which is significant and negative (the coefficient is significant; b=-0.446, p<.01). The result holds consistently in Models 3, 4 and 5 with the coefficient equalling -0.396, p<.05 in Model 5. This variable does indeed point to the negative effect of task complexity on timely completion supporting Hypothesis 1. Moreover, analyzing the change in probabilities for timely completion while keeping other variables at their mean, a one standard deviation change in task complexity increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 7.3% while it decreases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 9.7%.

Hypothesis 2: *Task Breadth.* I predicted that the breadth of tasks involved in a clinical trial represented by the number of conditions addressed in a clinical trial will have a negative effect on the timely completion of the trial. In Model 3, the coefficient for task breadth variable was statistically significant. The main effect of the task breadth variable was negative (b=-0.338, p<.01), this effect remains significant throughout the rest of the models, and in the final model as well (b=-0.362, p<.01). This result supports my hypothesis: there are costs to having a wider breadth of tasks involved in a clinical trial process; the effect of number of conditions addressed on a clinical trial on the timely completion of the trial is negative. Looking at the changes in probabilities for timely completion, it is possible to see that while keeping other variables at their mean, a one standard deviation change in task breadth increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 9.8% while it decreases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 13.2%.

Hypothesis 3: *Task Applicability.* I argued that the applicability of tasks as represented by the age range of subject pool desired for the study will have a positive effect on the timely completion of the study. Given Model 4's improvement over prior models, I inspect the coefficient of task applicability variable which is significant and positive (the coefficient is significant; b=0.0187, p<.05). The result holds consistently with the coefficient equalling 0.0161, p<.1 in Model 5. Given this positive and significant result, I find support for Hypothesis 3. Analyzing the change in probabilities for timely completion while keeping other variables at their mean, a one standard deviation change in task applicability decreases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 4.4% while it increases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 5.9%.

Hypothesis 4: *Partner Complexity.* I argued that the partner complexity and the associated costs of coordination as captured by the number of sponsors on a given trial will negatively effect the timely completion of the clinical trial. In Model 5, the coefficient for partner complexity variable was statistically significant. The main effect of the task breadth variable was negative (b=-0.545, p<.01). This result lends support to my hypothesis: there are coordination costs related to having more sponsors involved in clinical trial process which reflect themselves negatively on the timely completion of the clinical trial. Moreover, analyzing the change in probabilities for timely completion while keeping other variables at their mean, a one standard deviation change in partner complexity increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same

phase) by 11.6% while it decreases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 15.4%.

3.5 Discussion

The design of distributed work projects is a relatively understudied topic that not only involves modularization and task decomposition but also includes task selection and complexity. There are important coordination mechanisms such as learning, capability development, feedback and communication that assist in coordinating task and partner complexity in distributed work; however the task characteristics and how they influence the coordination of distributed work projects is not well known. Even though recent studies systematically document that geographical dispersion, team familiarity and task complexity effect the performance of distributed work in software development (Espinosa *et al.*, 2007) as well as patent examination process (Harhoff *et al.*, 2009) and contractual alliances (White *et al.*, 2005), there's little work that specifically disentangles the effect of task characteristics on the effective coordination and performance of distributed work projects. The primary issue of concern in this study is to disentangle the effects of task and partner complexity, task breadth and task applicability on the performance of distributed work projects. Hence, in this study I document how different task attributes create complexities which can hinder the performance of a project and cause potential losses in terms of time and money. Below, I explain the implications of these findings.

First, I supply empirical evidence for the effects of task complexity and its attendant coordination problems on project performance in clinical drug trials in diabetes. My primary contribution lies within the emphasis of complexity of clinical trial tasks in terms of task size, procedural complexity and coordinative complexity that stems from interdependencies among the processes involved (Campbell, 1988; Espinosa *et al.*, 2007; Wood, 1986). A number of

researchers have documented that structural complexity and work distribution can have negative effects on project performance (Bensaou *et al.*, 1995; Campbell, 1988; Espinosa *et al.*, 2007; Wood, 1986). A much less attended issue that I attempt to bring to the fore is the coordinative complexity stemming from interdependencies of tasks when multiple tasks of similar and interactive nature are taking place, as can be observed in clinical trials involving multiple interventions. By analyzing the impact of task complexity on project performance, one can reach a more fine-grained understanding of the design of distributed work projects that give the optimum project performance.

Second, I was able to analyze the effects of task breadth on the outcomes of distributed work in clinical trials. Having a wide arrange of conditions targeted in a clinical trial increases the costs of coordination as it increases the complexity and difficulty of communication, task allocation and effective management of distributed work projects (Huckman *et al.*, 2008). Increase in breadth widens task dimensions that require attention and at the same leads to higher complexity in understanding relationships across treatments, hampering project performance.

Third, applicability of the task in diverse settings was expected to have a positive impact on the timely completion of the project. Task applicability is related to the level of routinization of the underlying processes involved in the task. Tasks that are to be applied to a greater variability of settings are more easily routinizated and show efficiency benefits in implementation.

Finally, I find evidence for negative effects of partner complexity on project outcomes in terms of the timely completion of projects. My findings build up on previous literature and further tease out the effects of coordination problems that hamper collaborations in joint work efforts (Cummings *et al.*, 2005, 2007; Gulati *et al.*, 1998; Kiesler *et al.*, 2002). Attending to such

problems is important as the number of collaborations in sponsoring has increased steadily over the last decade, driven by several different motives including the need to collaborate and innovate in a rapidly developing field (Powell *et al.*, 1996; Rothaermel *et al.*, 2006; Rothaermel *et al.*, 2007). Given the results of this study, the partner complexity and associated coordination costs entail an issue of concern for organizations willing to jointly undertake distributed work projects.

3.5.1 Limitations

In spite of this study's contributions, limitations exist. I follow a reduced form approach to investigate the effects of design choices in terms of task and partner complexity, task breadth and applicability on the actual outcomes rather measuring directly intermediate processes and mechanisms. In this study, I measured task complexity, breadth and applicability related to clinical trial tasks on the outcome of the clinical trial process, but due to data limitations, I could not have intermediate level data to actually measure the duration of task routines and problems that arise during clinical trial processes. Also, I do not have a control for the clinical site activity outside the diabetes research which may be influencing the outcomes of clinical trials in the dataset. Inclusion of further data on other trials would help solve this problem.

3.6 Conclusion

Taken as a whole, the results provide evidence that design choices and selection of task attributes in terms of task complexity, breadth and applicability as well as partner complexity affect the project performance in distributed work and hence organizations designing and implementing such projects should be aware of the consequences of the design choices that they make. More importantly, although much attention surrounds collaborative work, an issue that has been somewhat neglected is the coordination costs that constitute the downside of such

collaboration. This study disentangles the effects of task characteristics on the speed to completion of projects that are distributed across organizations. Further research should focus on the dynamic processes that are involved in the implementation of task characteristics by looking at intermediate level data that pertains to underlying mechanisms. Doing so will bring a better understanding of not just the effect of project design and task characteristics on project outcomes as was tackled in this paper, but further our understanding of how experience in certain types of tasks may give rise to an ability to effectively coordinate distributed work projects.

Overall, this study contributes to the theories of organizational design, complexity and coordination of distributed work projects. It helps disentangle the relationship between characteristics of tasks in terms their breadth, applicability and complexity and the outcomes of distributed work. The study highlights the potential problems involved with tasks that involve processes which are interdependent and which may interact to curb project performance such as those in clinical trials involving multiple interventions.

CHAPTER 4

The Design and Management of Distributed Work

4.1 Introduction

Scholars have suggested that organizations can develop an ability to effectively manage inter-organizational relationships by having greater experience in managing such relationships (Anand *et al.*, 2000; Hoang *et al.*, 2005; Schreiner *et al.*, 2009; Zollo *et al.*, 2002). Moreover, coordination by planning and design of joint work have been shown as effective ways of managing inter-organizational activity (Hoang *et al.*, 2005; Kale *et al.*, 2002; Schreiner *et al.*, 2009). At the same time, recent work has shown that organizations can implement practices to capture, codify and internalize relevant know-how that enhances the organizations' ability to effectively manage inter-organizational joint work (Kale *et al.*, 2007).

Research in inter-organizational joint work points to factors affecting the success of such collaborations based on the organizations' ability to select partners that have the right complementarities and fit (Dyer *et al.*, 1998; Hitt *et al.*, 2000) and the right organizational structures (Gulati *et al.*, 1998; Hennart *et al.*, 2005; Schreiner *et al.*, 2009). At the same time the success of collaborative work is also shown to be related to the ability of organizations to manage such work in terms of task coordination, conflict resolution and knowledge sharing (Doz, 1996; Kumar *et al.*, 1998; Madhok *et al.*, 1998; Van de Ven *et al.*, 1992).

Inter-organizational collaborations face coordination challenges that stem from interdependence between partners (Gerwin, 2004; Gulati *et al.*, 1998; Litwak *et al.*, 1962) that can partly be addressed at partner selection and design stages (Gulati *et al.*, 1998). However, not all coordination challenges can be solved before hand, therefore mechanisms that help build the ability of organizations to effectively manage their joint work are of great importance.

The importance of coordination in collaborations is paramount given the fact that partners need to specify roles and responsibilities based on task characteristics and adapt to changes (Gerwin, 2004; Gulati *et al.*, 1998; Lawrence *et al.*, 1967; Litwak *et al.*, 1962; Schreiner *et al.*, 2009). This means that the ability to coordinate stems from knowledge and skills to identify the interdependence and manage it as the need for the adaptation arises. Although routine tasks may be specified easily, more complex and non-routine tasks require higher levels coordination ability.

Although there has been a plethora of research dealing with diverse aspects of distributed work (Cummings *et al.*, 2005, 2007; Hinds *et al.*, 2002; Hinds *et al.*, 2005; Srikanth, 2010), the effect of project design complexity on coordination of the distributed work hasn't received as much attention as would be expected. And although previous research has pointed to the importance of prior experience in inter-organizational relations as a way to develop an ability to effectively manage inter-organizational relationships (Anand *et al.*, 2000; Hoang *et al.*, 2005; Zollo *et al.*, 2002), little has been explored in terms how complexity and knowledge relevance affect the build up of coordination ability.

Given that scarce attention has been paid to the differential effects of joint work experiences in enhancing the ability to overcome coordination problems in distributed work projects, the primary issue of concern in this study is to disentangle the effects of complexity stemming form project design and partner selection while showing the effects of prior joint work experience in resolution of coordination problems. This study particularly analyzes the experience enhancing effects of knowledge relevance and task complexity in distributed work projects.

4.2 Theory Development and Hypotheses:

4.2.1 Design, Complexity and Coordination:

In designing distributed work projects, one issue that designers have to face is determining the optimal level of complexity for project performance. The projects' complexity arises from several factors including the size as well as the difficulty of tasks, the necessity to coordinate action across project partners and across locations.

The tasks that are included in a distributed work projects determine the extent to which the degree of interdependence varies across parties. In settings where complex and overlapping labor requires continuous adjustments between partners and force them to link specific activities with each other closely and regularly, a high degree of interdependence exists (Gulati *et al.*, 1998). For parties involved in joint work, the anticipated level of interdependence and coordination costs increase as the need for ongoing task coordination and joint decision making between them gains importance (Gulati *et al.*, 1998).

In the clinical trial setting, the involvement of sponsor organizations creates a need to coordinate responses and make adjustments to the trial processes on an ongoing basis as they receive information from the trial sites with regards to the performance of proposed treatments. The trial process may give rise to unexpected events such as adverse reactions that require adjustments to the protocol (Azoulay, 2004). As these adjustments require joint decision making given the stakes of sponsors in clinical development, the coordination costs increase. The involvement of multiple sponsors increases the complexity and difficulty of coordination (Cummings *et al.*, 2007; Hoang *et al.*, 2005) which in turn decreases the reaction time and slows down the clinical trial process.

In clinical trials, the timely completion of the distributed work across multiple clinical trials sites requires the effective coordination and management of these tasks across diverse locations. The extent of dispersion as seen by the increase in the number of locations where the tasks are taking place, adds an increased dimension to the problem of coordination and knowledge sharing in distributed work environment. Dispersion of the distributed work project increases temporal, spatial and demographic distance and reduces the ability of joint problem solving, communication and coordination all of which result in increased complexity and reduced performance in terms of timely completion.

One of the factors that influence the complexity of the project is the size of the task (Espinosa *et al.*, 2007). Size in terms of the number of instructions has been shown as an important factor determining the complexity of a task in settings such as software development (Banker *et al.*, 1998; Espinosa *et al.*, 2007) and patent examination (Harhoff *et al.*, 2009). At the same time, procedural complexity increases when projects require simultaneous implementation of multiple actions. In the clinical trial setting, interventions which include use of drugs, genes, vaccines and devices as ways of intervening with the state of health of volunteers under study (FDA, 2010) can be deemed an attribute that relate to the complexity of project design. As the number of interventions in a clinical study increases, both the complexity stemming from task size as well as the complexity stemming from involvement of complicated procedures surges. In other words design complexity may not only increase with task size but also with the amount of information and number and types of interrelationships between sub-processes related to these (Espinosa *et al.*, 2007) tasks. This is similar to the concept of coordinative complexity discussed in literature (Wood, 1986). The increased levels of complexity in design will result in difficulties

in coordinating distributed work projects and negatively impact the performance of such projects.

Hence, I propose that;

Hypothesis 1a: The effect of the number of sponsors in a clinical trial on the timely completion of the clinical trial is negative.
Hypothesis 1b: The effect of the number of interventions that are involved in a clinical trial on the timely completion of the clinical trial is negative.
Hypothesis 1c: The effect of the number of unique locations where clinical trial sites are located in a given clinical trial, on the timely completion of the clinical trial is negative.

4.2.2 Organizational Learning and Relational Capital

Collaborative linkages among organizations can be associated with two distinct kinds of benefits. They can provide the benefit of resource sharing, allowing researchers to combine knowledge, skills, and access to physical assets (e.g. labs). Collaborative linkages also provide access to knowledge spillovers, serving as information conduits through which news of technical breakthroughs, new insights regarding problems, or failed approaches travels from one entity to another. Another important aspect of collaborations is that collaborating organizations develop an ability to coordinate action and manage joint work through gaining experience in working with one another.

Individually and organizationally held knowledge has been attributed to be a basis for creating firm-level capabilities that act as a source of competitive advantage (Grant, 1996). The development of capabilities has been referred to as the process by which organizations generate new routines or modify existing routines and resource configurations to create a potential to achieve particular outputs that are congruent with organizational goals (George *et al.*, 2008). Within the capabilities literature a stream of researchers has focused on the inter-organizational relationships and pointed to the existence of inter-firm resources and routines which span boundaries of the organizations and support the organizations' competitive positions (Dyer *et al.*,

1998). In particular, relational capital has been associated with firm's ability to create relational rents from its partners. As a specific type of relational capital, alliance management capability was also shown to affect the performance of alliances in terms of both financial and managerial outcomes (Kale *et al.*, 2002; Kale *et al.*, 2007). Hence, the deployment of capabilities in organizing joint action and collaborative work has been related to positive joint performance outcomes.

Scholars have shown that experience results in the development of capabilities that facilitate coordination (Autio *et al.*, 2008; Dyer *et al.*, 1998; Hoang *et al.*, 2005; Kale *et al.*, 2002; Kale *et al.*, 2007). Improved knowledge codification and enhanced cooperation over collaborative activities results in more effective coordination. Firms with prior experience in collaborations perform better than those without such experience (Hoang *et al.*, 2005). As a benefit of experience in collaborative work, organizations learn to coordinate action and collaborative experiences in alliances positively affect performance (Shan *et al.*, 1994) and create value (Anand *et al.*, 2000).

One of the factors that influences learning in collaborative work is the context in which learning takes place (Hoang *et al.*, 2005; Kale *et al.*, 2007). Through prior collaborative experiences in a given knowledge domain, collaborators can gain a profound understanding of the key problem areas and how their respective knowledge stocks complement one another. At the same time, the types of problems that arise in joint work projects are related to the knowledge area on which the projects are building. Hence working together in a given knowledge area allows organizations to learn how to effectively overcome further problems that may arise in the same domain in future collaborations.

As organizations learn to coordinate their activities in a given setting, i.e., subject area, they learn specific skills with regards to the knowledge domain which allows them to reap the benefits of their joint work at a more fine-grained level. Moreover, as the level of complexity increases, the coordination of activities becomes more difficult, requiring higher intensity of on-going communication and information exchange. Given the positive role of experience in coordination of knowledge-intensive and complex tasks, I hypothesize that:

> *Hyptothesis2a. The effect of the number of times that a sponsor and a clinical trial site have worked together in all clinical trials, on the timely completion of the focal clinical trial is positive.*
> *Hyptothesis2b. The effect of the number of times a sponsor and a clinical trial site have worked together in clinical trials in the same knowledge domain as the focal trial, on the timely completion of the focal trial is positive.*
> *Hyptothesis2c. The effect of the number of times a sponsor and a clinical trial site have worked together in clinical trials in the same knowledge domain as the focal trial and involving multiple interventions, on the timely completion of the focal clinical trial is positive.*

4.2.3 Learning and Task Complexity:

As previous studies have shown, experience has a critical role in the build of capabilities that allow organizations to effectively coordinate action and manage collaborative work (Dyer *et al.*, 1998; Hoang *et al.*, 2005; Kale *et al.*, 2007). Researchers have analyzed the various types of experience in inter-organizational joint work as they pertain to exploration and exploitation frameworks (Gulati *et al.*, 2009; Hoang *et al.*). Although the classification of experience stemming from knowledge context and work design is critical in understanding the build up of coordination ability, not much attention has been paid to this branch.

The context of learning is an important factor determining the extent to which experience can be effectively leveraged in further work. Collaborating organizations learn knowledge specific attributes of problems that arise in their knowledge setting. When they need to coordinate inter-organizational joint work in the same knowledge domain in future work, the

organizations can fall back on the learnt routines in problems solving, sense-making and communicating in projects using this knowledge base. Hence, experience stemming from projects in the same knowledge domain as the focal project is more effective then experience from joint work in other knowledge domains.

Hypothesis 3a. The effect of joint experience on the performance of a clinical trial is higher when the prior experience stems from the same knowledge domain as the focal experience.

The intensity of the use of coordination mechanisms, determines the speed with which an ability to coordinate action can be built across two collaborating organizations. As the intensity of mechanism usage increases, one can expect to see higher levels of coordination ability building up through experience. Hence, organizations that work together on projects involving highly complex designs can build a higher ability to coordinate action. As complexity of the joint work project increases, the demand on communication and modifications involving changes in plans and adaptations to changing conditions increases. For organizations involved in joint work, developing inter-organizational capabilities to work in complex tasks becomes an important factor enhancing their ability to undertake complex projects. And, for organizations that have worked together in projects involving complex tasks, the experience they gain through coordinating such complex tasks requiring extensive use of coordination mechanisms is even more valuable then the experience gained through simple projects that don't require intense coordination.

Hence, I propose that;

Hypothesis 3b. The effect of joint experience from a given knowledge domain on the performance of a clinical trial is higher when the prior experience from a given knowledge domain involves complex tasks and when the focal project in the given knowledge domain involves complex tasks.

4.3 Method

4.3.1 Sample and data

To examine the effects experience on the resolution of coordination problems and to tease out the differential effects of different types of experience on project performance in distributed work projects, I use clinical trials as a setting. Firms that wish to gain regulatory approval for market introduction of their drugs need to provide substantial evidence of their drug's effectiveness through controlled clinical trials (FDA, 2010). Clinical trials involve critical tasks such as recruitment of candidates, following up on the protocol set for study, submission of case reports including original patient records and charts (Azoulay, 2004). Clinical trials offer an ideal setting to study the coordination of distributed work projects because clinical trials involve drug testing projects that are distributed across a variety of trial locations and that follow tasks as defined by the main protocol. Tasks involved in clinical trials vary by phase of the drug development process, the conditions addressed, the interventions at use and other trial specifications that determine the routine manipulation, storage and sharing of symbolic information within established categories (Azoulay, 2004; Huckman *et al.*, 2008). At the same time, clinical trials also involve knowledge production which occurs through generation of search rules for identifying problems and heuristics that leads to their solution and which can't be reduced in simple protocol steps (Azoulay, 2004). In this sense, each clinical trial displays a unique set of characteristics which reflects itself on the complexities of tasks involved in clinical research and the interdependencies between these tasks' attributes.

The sample to test the hypotheses consists of 4769 clinical trial observations from 5 disease areas including diabetes, arthritis, liver disease, cardio diseases and asthma, between the years of 1981 and 2009. I have retrieved clinical trial data from the Federal Drug

Administration's website that includes data on all clinical trials associated with treatments seeking regulatory approval in the United States. Although the application for approval is made in the United States, this does not necessarily mean clinical trials have to or do exclusively take place in the US. In fact, increasingly the trials are distributed on a global scale. However, since the United States is the main market for pharmaceutical companies, the FDA database which is the most comprehensive of its sort covers a significant portion of the universe of all clinical trial studies carried across the globe. I have retrieved clinical study data by using relevant search words for identifying studies that are addressing the relevant conditions and downloaded all available information. After downloading the data, I have gone through 19800 clinical trial sites associated with the 4769 trials listed in the dataset to correct for potential mistypes. I have also matched clinical research organizations' affiliations by using the AAPS News Magazine Buyer's Guide, company directories to control for potential differences stemming from the involvement of clinical research organizations.

4.3.2 Measures

Dependent Variable.

Timely Completion. To assess the impact of design complexity and experience based coordination ability on the performance of clinical trials, I have created a timely completion variable that calculates the distance of same class, treatment, starting year based trials in terms of standard deviation(s) from the mean completion time of all similar studies in the dataset. This measure is calculated by subtracting the mean time of completion from similar trials from the completion time of the focal trial and dividing this value by the standard deviation of completion time of trials from similar trials' pool. The resulting value is multiplied by -1 for ease of

interpretation; hence the timely completion value increases as the time to completion of a trial with respect to similar trials' performance improves.

In order to overcome the potential problem of mismatches in creating these categories, I have created respective categorical scales for the three phases that are observed in the dataset. For example, Phase I studies are "conducted to determine the metabolism and pharmacologic actions of drugs in humans, the side effects associated with increasing doses, and to gain early evidence of effectiveness" (FDA, 2010) and deal with initial efficacy and safety checks which take significantly less time to complete. Therefore each Phase I study's completion time is compared against the respective mean completion time and standard deviation distances of the population of Phase I studies rather than the measures for the entire set of clinical trials. The same procedure is applied for constructing the respective categories of Phase II and Phase III studies. Timely completion is an important measure of clinical trial success since sponsors of the clinical trial depend on FDA approval for marketing their drugs (Azoulay, 2004) and potential delays may cause losses in terms of monetary benefits and may even harm the reputation and market position of sponsoring firms.

Independent Variables

Number of Interventions. This measure calculates the number of interventions listed on a clinical trial. It captures the complexity of the clinical trial given the number of interventions listed in the trial design.

Number of Study Designs. This measure calculates the number of study designs listed on a clinical trial. This construct is used to capture the complexity of clinical trial.

Number of Locations. This measure records the number of locations in which the clinical trial is conducted.

Number of Sponsors. This measure records the number of sponsors that a clinical trial has. This construct is used to capture the complexity of the trial stemming from having to coordinate trial across multiple partners.

Lead sponsor's Experience Overall. I calculate the number of clinical trials that the lead sponsor was engaged in before starting the focal clinical trial.

Lead sponsor and Locations' Joint Experience Overall. This variable calculates the number of times that a lead sponsor has worked together with a location (clinical trial site) before the focal trial.

Lead sponsor and Locations' Joint Experience in a Given Subject Area. This variable calculates the number of times that a lead sponsor has worked together with a location in a given subject (disease area) prior to the focal clinical trial.

Lead sponsor and Locations' Joint Experience in Interventions in a Given Subject Area. This variable calculates the number of interventions on which a lead sponsor has worked together with a location in a given subject area prior to the focal clinical trial. It calculates the aggregate number of interventions implemented in previous clinical trials in the subject area of focal clinical trial, where the lead sponsor and the location were jointly involved.

Control Variables

Trial Start Year. This variable records the year in which the clinical trial was initiated.

Enrollment. This variable records the number of patients that are enrolled for the focal trial.

Healthy Volunteers. This is a dummy variable that takes the value of 1 if healthy volunteers could be included in the trial and 0 if healthy volunteers can not take part in the focal clinical trial.

Age Range. This variable controls for the trial. Depending on the trial requirements, age range varies across clinical trials where participants from different age groups may be sought. The smallest age for participation is designated as 11 in this sample and 73 is the oldest participant in the range.

Percentage of In-house Locations. This variable controls for the intensity of in-house sites in the clinical trial process. The number of in-house (those trial sites that belong to the lead sponsor) sites is divided by the number of overall clinical trial sites for a given clinical trial.

Industry Dummy. This measure takes the value of 1 if there are industry sponsors on the clinical trial and 0 otherwise.

4.3.3 Analysis

The dependent variable in this paper is the timely completion of clinical trials as calculated from their relative performance to clinical trials started in the same year within the same disease area and focusing on the same phase of trials. This allows for a direct comparison between the performances of clinical trials across large sets of groups. I create the "similar clinical trial pools" for comparing performance. I employ ordinary least squares regressions including controls for lead sponsor, eligibility and sponsor type.

4.4 Results

In Table 6, the descriptive statistics and correlations of the variables are reported. Although the correlations reported are relatively low, I derived the variance inflation factors (VIFs) as a control from the final model and the average VIFs were well below the general accepted threshold of 10 (Greene, 1997). VIFs are typically made as a post estimation command for linear regressions, but as they make an assumption about the relationship between the independent variables, it is possible to use this technique for other functional forms as well

(Menard, 2002). In testing the hypotheses, I also included the variables stepwise to check and make sure that the signs of coefficients are stable across the regressions. If multicollinearity would have been a major issue, sign and coefficients could have changed direction. Taken together, these precautions minimized the problem of multicollinearity.

TABLE 6: DESCRIPTIVE STATISTICS AND CORRELATIONS

	Variables	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
1	Timely Completion	.00	1.00	-10.27	1.85	1														
2	Trial Start Year	2005.11	3.79	1965	2009	.65	1													
3	Enrolment	431.66	2952.77	1	228834	-.05	-.04	1												
4	Healthy Volunteers	.15	.36	0	1	.03	.04	.03	1											
5	Age range	56.84	13.70	11	73	-.03	-.08	.04	-.22	1										
6	Percentage of In-house Locations	.37	.46	0	1	-.03	-.04	0	.01	.06	1									
7	Industry Dummy	.35	.48	0	1	.22	.16	0	-.14	.02	-.31	1								
8	Number of Interventions	1.92	1.38	1	15	-.01	.08	0	.01	-.02	-.07	.10	1							
9	Number of Study Designs	6.99	3.19	1	12	.39	.60	-.02	.04	-.08	-.09	.17	.10	1						
10	Number of Locations	14.25	52.52	1	1367	.02	.03	.11	-.10	.06	-.12	.32	.09	.03	1					
11	Number of Sponsors	1.53	.98	1	14	-.07	-.03	.04	.07	-.03	.04	-.30	0	-.01	-.08	1				
12	Lead sponsor's Experience Overall	171.07	298.56	1	1685	.01	.10	0	.08	-.04	-.06	.07	.07	-.01	.06	-.13	1			
13	Lead sponsor and Locations' Joint Experience Overall	8.31	24.36	1	425	-.01	-.04	-.02	.01	.08	.26	-.12	-.05	-.17	-.05	-.09	.28	1		
14	Lead sponsor and Locations' Joint Experience in a Given Subject Area	2.73	7.50	1	132	-.03	-.09	-.02	-.01	.09	.19	-.10	-.04	-.20	-.04	-.09	.36	.58	1	
15	Lead sponsor and Locations' Joint Experience in Interventions in a Given Subject Area	5.45	21.17	1	456.44	.02	0	-.01	-.02	.08	.16	-.05	.21	-.10	0	-.09	.35	.57	.88	1

TABLE 7: OLS MODELS FOR TIMELY COMPLETION OF CLINICAL TRIALS

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6	Model 7
Trial Start Year	0.182***	0.208***	0.224***	0.201***	0.203***	0.202***	0.217***
	[0.00985]	[0.0130]	[0.0136]	[0.0132]	[0.0128]	[0.0129]	[0.0137]
Enrolment	-4.93e-05***	-3.69e-05***	-3.69e-05***	-3.67e-05***	-3.68e-05***	-3.73e-05***	-3.71e-05***
	[8.30e-06]	[8.22e-06]	[8.39e-06]	[8.21e-06]	[8.24e-06]	[8.27e-06]	[8.43e-06]
Healthy Volunteers	0.185***	0.193***	0.198***	0.196***	0.194***	0.198***	0.205***
	[0.0415]	[0.0479]	[0.0475]	[0.0476]	[0.0479]	[0.0479]	[0.0473]
Age range	0.00607***	0.00601***	0.00561***	0.00578***	0.00560***	0.00541***	0.00493***
	[0.00163]	[0.00190]	[0.00193]	[0.00189]	[0.00190]	[0.00188]	[0.00191]
Percentage of In-house Locations	0.0941**	0.0975**	0.0919**	0.0312	0.0363	0.0349	0.00479
	[0.0371]	[0.0421]	[0.0419]	[0.0447]	[0.0457]	[0.0454]	[0.0455]
Number of Interventions		-0.0601***	-0.0619***	-0.0565***	-0.0585***	-0.0721***	-0.0739***
		[0.0120]	[0.0119]	[0.0120]	[0.0120]	[0.0123]	[0.0140]
Number of Study Designs		-0.0140**	-0.00769	-0.0106	-0.00974	-0.0114*	-0.00371
		[0.00685]	[0.00699]	[0.00694]	[0.00680]	[0.00686]	[0.00682]
Number of Locations		-0.00116**	-0.00124**	-0.00111**	-0.00113**	-0.00112**	-0.00118**
		[0.000516]	[0.000554]	[0.000515]	[0.000519]	[0.000521]	[0.000564]
Number of Sponsors		-0.0311**	-0.0285*	-0.0286*	-0.0283*	-0.0282*	-0.0246
		[0.0158]	[0.0158]	[0.0157]	[0.0157]	[0.0157]	[0.0158]
Lead sponsor's Experience Overall			0.000579***				0.000655***
			[0.000218]				[0.000219]
Lead sponsor and Locations' Joint Experience Overall				0.00480***			0.00341**
				[0.00114]			[0.00141]
Lead sponsor and Locations' Joint Experience in a Given Subject Area					0.00916***		0.00466
					[0.00228]		[0.00569]
Lead sponsor and Locations' Joint Experience in Interventions in a Given Subject Area						0.0135***	0.0151**
						[0.00293]	[0.00696]
Log Likelihood	-4971	-4077	-4060	-4059	-4058	-4053	-4027
R-Squared	0.621	0.636	0.638	0.639	0.639	0.639	0.643

* p<.1, ** p<.05, ***p<.01; two-tailed tests. Robust standard errors are in brackets. 4769 clinical trial observations; Leadsponsor fixed effects are included in all models. Trial phase, Subject Eligibility and Sponsor Type controls are included in all models.

Table 7 shows the results of the OLS regression analysis where the dependent variable is the timely completion of clinical trials. Model 1 shows the baseline model which includes controls for clinical trial start year, industry dummy, age range and percentage of in-house locations. In Model 2, I introduce four of the theory variables of interest (number of interventions, study designs, locations and sponsor). Model 3 adds the terms for lead sponsor experience while Model 4 includes lead sponsor's joint experience with locations. Model 5 includes the lead sponsor and location joint experience in a given disease area while Model 6 includes the cumulative number of interventions undertaken by a lead sponsor and locations in a given subject area. Model 7 includes all variables of concern. Each model represents a significant improvement over the baseline model with the log likelihood value increasing from –4971 for the base model ($p<.001$) to –4027 ($p<.001$) for the final model represented in Model 7 of Table 7. The baseline model was generally consistent with my expectations regarding the effects of eligibility and start year.

Hypothesis 1a: *Number of Sponsors.* I predicted that the number of sponsors, which is defined as the number of sponsors funding a clinical trial will have a negative effect on the timely completion of the clinical trial. In Model 2, the coefficient for the number of sponsors variable is negative and significant as expected. ($b=-0.0311$, $p<.05$), this effect remains significant across the rest of models until Model 7. This result lends support to the hypothesis that number of sponsors is negatively related to the timely completion of clinical trials and is in agreement with my findings from chapter 3.

Hypothesis 1b: *Number of Interventions.* I predicted that the number of interventions, which is defined as the number of interventions implemented in a clinical trial protocol will have a negative effect on the timely completion of the clinical trial. In Model 2, the coefficient for the

number of interventions variable is negative and significant as expected. (b=-0.061, p<.01), this effect remains significant across the rest of models. This result lends support to the hypothesis that number of interventions is an accurate measure of the complexity of the clinical trial and is negatively related to the timely completion of clinical trials.

Hypothesis 1c: *Number of Locations.* I predicted that the number of locations, which is defined as the number of locations involved in a clinical trial will have a negative effect on the timely completion of the clinical trial. In Model 2, the coefficient for the number of locations variable is negative and significant as expected (b=-0.0011, p<.05), this effect remains significant across the rest of models. This result lends support to the hypothesis that number of locations measures a pier of complexity that negatively effects the timely completion of clinical trials.

Hypothesis 2a: *Lead Sponsor and Locations Experience Overall.* I argued that the joint experience of lead sponsors with locations is positively related to the timely completion of clinical trials. In Model 4, I include the lead sponsor and location experience overall variable which is significant and positive (b=0.0048, p<.01) and observe that this effect is shown in the final model as well. This lends support to the hypothesis that joint experience of lead sponsors and locations helps overcome coordination problems and effectively complete clinical trials.

Hypothesis 2b: *Lead Sponsor and Locations Experience in the Focal Trial Disease Area.* I argued that the joint experience of lead sponsors with locations from previous trials that are focusing on the same disease area as the focal trial is positively related to the timely completion focal trials. In Model 5, I include the lead sponsor and location experience in a given subject area variable which is significant and positive (b=0.0092, p<.01). This lends support to the hypothesis

that joint experience of lead sponsors and locations in a given subject area helps overcome coordination problems involved in clinical trials.

Hypothesis 2c: *Lead Sponsor and Locations Experience in Interventions in the Focal Trial Disease Area.* I argued that the joint experience of lead sponsors with locations in interventions from previous trials that are focusing on the same disease area as the focal trial is positively related to the timely completion focal trials. In Model 6, I include the lead sponsor and location experience in a given subject area variable which is significant and positive (b=0.0135, p<.01). This effect exists in the final model as well. This lends support to the hypothesis that joint experience of lead sponsors and locations in interventions in a given subject area helps overcome coordination problems involved in clinical trials.

Hypothesis 3a: *Lead Sponsor and Locations Experience in the Focal Trial Disease Area.* I argued that the joint experience of lead sponsors with locations from previous trials that are focusing on the same disease area as the focal trial is positively related to the timely completion focal trials. Moreover that this effect is more prominent than the effect of overall experience observed in Model 4. Indeed the difference between the effect sizes is almost double-fold, where as the experience stemming from a given disease area has a positive effect of 0.0048, the positive effect of experience from the focal disease area is at 0.0092 levels. Both measures are significant at the p<.01 level. This lends support to the hypothesis that the effectiveness of learning from joint work is enhanced when the joint work experience stems from the same area as the focal project area.

Hypothesis 3b: *Lead Sponsor and Locations Experience in Interventions in the Focal Trial Disease Area.* I argued that the joint experience of lead sponsors with locations in interventions from previous trials that are focusing on the same disease area as the focal trial is

positively related to the timely completion focal trials. Moreover that this effect is more prominent than the effect of only subject area specific experience observed in Model 5. I had already established in Hypothesis 1b that interventions create a level of complexity that slows down the clinical trials. Given this critical nature of interventions, experience in interventions in a given area is indeed more important than just experience in the given disease area. When the effect sizes are analyzed it can be seen that where as the experience stemming from a given disease area has a positive effect of 0.0092, the positive effect of experience from the focal disease area is at 0.014 levels. Both measures are significant at the $p<.01$ level. This lends support to the hypothesis that the effectiveness of learning from joint work in the same knowledge setting increases as the project complexity increases.

4.5 Discussion

The role of experience in inter-organizational settings is a well studied topic; however there are important caveats to studying experience as an overall measure rather than analyzing the different levels of experience and the underlying logic with which they impact performance of collaborative work. The primary issue is one of harnessing the benefits of experience (by knowledge specificity and task relevance) while measuring the impact of potential coordination problems (by including measures of sponsor, location and protocol specific complexities). This study documents how the benefits of different types of collaborative experience impact the effective management of clinical trial processes. Further, I find evidence that prior experience in joint work not only positively effects the clinical trial completion times but also the effectiveness of these experiences is enhanced when the ties' experience stems from same the knowledge domain as that of the current project, perhaps because more relevant tasks are involved in the

current project or because learning based on specialization is especially important in this complex setting. The findings of this study relate to three pillars of knowledge.

First, I supply empirical evidence for the effects of location, protocol and partner specific complexities and its attendant coordination problems on timely completion of clinical trials. My primary contribution lies within the emphasis of coordination costs that constitute an important barrier on the effectiveness and efficiency of joint work in undertaking complex clinical trials. Particularly I show the importance of design related complexity in hampering project performance, this effect originates from project size, task distribution and partner related complexities.

Second, I show the positive effects of joint work experiences on overcoming potential coordination issues in the clinical trial setting. Especially the prior joint work between lead sponsor and locations (clinical trial sites) has an important positive effect on the success of focal clinical trials. This effect is valid for not just the overall joint work but subject and protocol specific work as well.

Third, heterogeneity among types of experience is relevant when one thinks of the layers of experience that are more pronounced in relation to the knowledge specificity and task complexity. Although, overall experience in joint work has been shown to be an important factor in mitigating coordination problems, a more fine-grained look at the effect of experience across levels shows that as the knowledge specificity and task complexity of experience increases, the impact of the experience in mitigating potential coordination problems increases in parallel. This is an important finding for practitioners and researchers who would like to better understand the dynamics of joint work and effective work distribution.

4.5.1 Limitations

In spite of this study's contributions, limitations exist. I follow the tradition of learning curve research of studying the actual outcomes of learning and prior experience rather than measuring directly intermediate processes and mechanisms (Arrow, 1962; Yelle, 1979). In this study, I measured experience outcomes related to learning on the timely completion of clinical trials, but due to data limitations, I could only observe the relevant performance measures of experience with regards to the completion and not with regards to the amount of money invested or intermediate success and failures that projects experienced. Although the overall measure should be correlated with such factors and hence not be of concern.

4.6 Conclusion

Taken as a whole, my results provide evidence of the positive effects of joint work experiences on overcoming potential coordination issues in the distributed work projects. Especially the prior joint work between lead sponsor and locations (clinical trial sites) has an important positive effect on the success of focal clinical trials. This effect is valid for not just the overall joint work but subject specific work as well. While having experience is important, perhaps more important is the type of experience that the parties have together. And with regards to this point, I have been able to show that knowledge specificity is an important attribute that determines the effectiveness of experience in that prior experience from same knowledge domains has a higher impact on the success of the focal project with respect to the overall experience. Moreover, task complexity is another characteristic of experience that determines the effectiveness of prior experiences on mitigating coordination problems in the current setting. The

higher the complexity of previous joint work projects, the more effective is the experience gained from such projects in mitigating coordination problems.

CHAPTER 5

Internationalization and Coordination of Distributed Work

5.1 Introduction

A stream of research that is increasingly receiving attention in the international business literature is that of the management of international R&D activities (Feinberg *et al.*, 2004; Gupta *et al.*, 1996; Mendez, 2003). Within this domain researchers have been interested in analyzing the factors determining international growth and knowledge management (Autio *et al.*, 2000; Bresman *et al.*, 1999; Minbaeva *et al.*, 2003). Others have looked at the effects of inertia on location and collocation in R&D networks (Narula, 2002; Narula *et al.*, 2009) and still others have analyzed factors influencing learning and knowledge transfer in international joint ventures (Barkema *et al.*, 1997; Dhanaraj *et al.*, 2004).

However, while international business literature has focused densely on the issues of knowledge transfer, little attention has been paid to the coordination of distributed work across boundaries (Kumar *et al.*, 2008; Srikanth, 2010). Prior work in this area has been dealing with two approaches to coordinating internationally distributed work: a) coordination through planning and b) coordination through feedback.

Coordination through planning mainly deals with task decomposition and modularization in order to avoid the necessity of on-going communication for coordinating action (Galbraith, 1977; March *et al.*, 1958; Tushman *et al.*, 1978). This approach, although feasible in large scale projects where modularization is possible and maintainers with profound knowledge of the system are readily available, may not be an option in smaller scale projects where upfront investments in modularization may not be viable (Srikanth, 2010). Therefore organizations may fall back on the alternative option of coordinating through feedback.

Coordination through feedback refers to on-going communication to overcome conflicts that stem from interdependencies (Hinds *et al.*, 2005). This type of coordination has been associated with collocation as a way to enable communication. The research in this area has focused on the benefits of collocation in terms of sharing a social context and being able to have face-to-face communication (Kiesler *et al.*, 2002). The importance of having shared artefacts, shared conventions and being aware of each others' work has also been emphasized as essential benefits of proximity in improving coordination (Cramton, 2001; Kraut *et al.*, 2002; Olson *et al.*, 2002). Spatial proximity is also shown to increase spontaneous communication and improve ease of knowledge sharing (Kraut *et al.*, 2002). The essence of these coordination mechanisms is similar to what is achieved by common ground through mutual knowledge of role and responsibilities (Bechky, 2006; Faraj *et al.*, 2006), the environment (Bechky, 2006; Clark, 1996) or background scientific knowledge (Puranam *et al.*, 2009).

Furthermore, a growing line of research focuses on tacit coordination mechanisms in off-shore development projects that deal with coordinating action through procedural, contextual and interpersonal common ground. Tacit coordination mechanisms build up on the above mentioned three common ground schemes to overcome coordination problems through neither planning, nor feedback but by using procedural, contextual and interpersonal common ground (Orlikowski, 2002; Srikanth, 2010).

Although, prior research has focused on these various coordination mechanisms to effectively manage internationally distributed work projects, no empirical research focused on whether and to what extent collocation (coordination by feedback) is important in settings where coordination by planning and modularity is expected. Hence, in this study I'll be exploring the extent to which complexities generated by the extent of internationalization, partner complexity

and collocation at country level, affect management and coordination of distributed work projects.

5.2 Theory and Hypotheses

5.2.1 Partner Complexity and Coordination

Coordination costs include anticipated organizational complexity of decomposing tasks among partners, the ongoing coordination of activities to be completed jointly or individually across organizational boundaries as well as the communication and decisions that would be necessary to carry out such activities (Gulati *et al.*, 1998). In settings where complex and overlapping labor requires continuous adjustments between partners and forces them to link specific activities with each other closely and regularly, a high degree of interdependence exists (Gulati *et al.*, 1998).

Clinical trials are sponsored by organizations which have to coordinate their responses and make adjustments to the trial process on an ongoing basis as they receive information from the trial sites with regards to the performance of proposed treatments. The trial process may be interrupted by unexpected events such as adverse reactions that require adjustments to the protocol (Azoulay, 2004). As these adjustments require joint decision making given the stakes of sponsors in clinical development, coordination of action becomes an important issue. The involvement of multiple sponsors increases the complexity in managing the trial process and the associated difficulty of coordination by increasing difficulty of joint decision making (Cummings *et al.*, 2007; Hoang *et al.*, 2005) which in turn decreases the reaction time adversely affecting the performance of clinical trial.

Therefore I posit that:

Hypothesis 1: The effect of the number of sponsors in a clinical trial on the timely completion of the clinical trial is negative.

5.2.2 Extent of Internationalization, Coordination and Knowledge Creation:

Coordination of distributed work in knowledge generation across geographical space is a time and effort consuming endeavor requiring high degrees of communication and synchronization of action across organizations (Cummings *et al.*, 2005, 2007).

The spatial dispersion (Kiesler *et al.*, 2002; O'Leary *et al.*, 2007) of organizations is an important factor in determining the efficiency of communication and the ability of workers to coordinate action (Hinds *et al.*, 2002). As it reduces the chances for face-to-face interaction, spatial dispersion leads to a decrease in spontaneous communication that would facilitate knowledge sharing and problem solving activities (Burke *et al.*, 1999; Dennis *et al.*, 1988). Geographic dispersion also brings about a temporal dimension which reduces the ability of real-time problem solving between organizations (Malone *et al.*, 1994; O'Leary *et al.*, 2007). To give an example, a clinical investigation team in Brazil may have conversations with their sponsors in the East coast of United States during regular office hours since they belong to the same time zone, however, this wouldn't be possible for the same team to communicate with their colleagues in China. Moreover, shared demographic characteristics such as shared culture, identity and context have been shown to moderate the potential coordination problems and conflicts that may arise during distributed work processes effecting efficiency, quality, technical innovation and adherence to procedures (Hinds *et al.*, 2005). In a similar nature, literature has also shown that similarity facilitates knowledge sharing in multinationals and that employee ability increases knowledge transfer (Lauring, 2009; Makela *et al.*, 2007; Minbaeva *et al.*, 2003).

In clinical trials, the timely completion of the distributed work across multiple clinical trial sites requires the effective coordination and management of tasks across diverse locations.

The extent of internationalization as seen by the increase in the number of countries where joint work in distributed projects is taking place, adds an increased dimension to the problem of coordination and knowledge sharing in the distributed work environment. Internationalization of the distributed work projects increases temporal, spatial and demographic distance and reduces the ability of joint problem solving, communication and coordination all of which hamper performance.

Hence, I propose that;

Hypothesis 2: The effect of the number of unique countries where clinical trial sites are located in a given clinical trial, on the timely completion of the clinical trial is negative.

5.2.3 Site Coordination Complexity:

The complexity of coordination in distributed work increases not only by geographical dispersion but also by the increase in interdependencies as can be seen through the number of organizations that carry out the distributed work. As the number of sites, which are distinct locations where drug testing occurs, increases so does the interdependencies which must be managed by the oversight of coordinating units. Although the tasks may be parallel, the advancement of a distributed work process may necessitate at least the accomplishment of certain objectives before different units can proceed to consequent steps (Azoulay, 2004; Hinds *et al.*, 2002; O'Leary *et al.*, 2007; Sarbaugh-Thompson *et al.*, 1998). For example, in the clinical trial setting, clinical trials cannot proceed without having fulfilled their recruitment quotas and this is determined independent of the ability of every single site to meet the criteria.

In the clinical trial setting, inclusion of multiple sites creates complexities in managing the trial processes where both data manipulation including storage and analysis of trial outcomes but more importantly knowledge generation in terms of generating search heuristics and

solutions to potential setbacks in trial process occur (Azoulay, 2004). The increased interdependencies arising from inclusion of multiple locations may hamper the performance of distributed projects. Therefore, I hypothesize that,

> *Hypothesis3: The effect of the number of clinical trial locations on the timely completion of the clinical trial is negative.*

5.2.4 Same country sites; Effects of Country Level Collocation in Distributed Work:

The complexity of coordination in distributed work increases by geographical dispersion and especially so in the case of international joint work since temporal, spatial and demographic barriers arise on the way of effective communication, joint decision making and knowledge sharing. Literature shows the importance of facilitators of knowledge transfer that include visits, and meetings, time elapsed and communication (Bresman *et al.*, 1999) which are more likely to take place when temporal and spatial gaps aren't highly pronounced.

The coordination and control of distributed project work across borders requires exchanges between geographically dispersed people of different cultural backgrounds (Mendez, 2003). The cultural differences may hinder communication and trust and disturb coordination of activities in task group (Watson *et al.*, 1993). A factor that would mitigate this challenge in coordinating distributed work across the globe would be the inclusion of hybrid mode internationalization including both local and global sites on a work project (Gassmann *et al.*, 1999). The complexities associated with spatial, temporal and cultural distance on coordination of activities will be reduced when the overseeing and the undertaking organizations share the same country background. Spatial, temporal and cultural proximity eases communication and fosters collaboration across sites and sponsors, leading to an increased performance of the distributed work projects.

Hence, I hypothesize that;

Hypothesis 4: The effect of the number of same country trial locations, on the timely completion of the clinical trial is positive.

5.3 Method

5.3.1 Sample and data

To examine the effects of internationalization and coordination complexity and to analyze how collocation as a coordination mechanism affects project performance in a distributed work environment, I use clinical trials as a setting. Firms that wish to gain regulatory approval for market introduction of their drugs need to provide substantial evidence of their drugs' effectiveness through controlled clinical trials (FDA, 2010). Clinical trials involve critical tasks such as recruitment of candidates, following up on the protocol set for study, submission of case reports including original patient records and charts (Azoulay, 2004). Clinical trials offer an ideal setting to study task complexity and coordination mechanisms because tasks included in clinical trials vary by phase of the drug development process, the conditions addressed, the interventions at use and other trial specifications that determine the routine manipulation, storage and sharing of symbolic information within established categories (Azoulay, 2004; Huckman *et al.*, 2008). At the same time, clinical trials also involve knowledge production which occurs through generation of search rules for identifying problems and heuristics that leads to their solution and which can't be reduced in simple protocol steps (Azoulay, 2004). In this sense, each clinical trial displays a unique set of characteristics which reflects itself on the complexities of tasks involved in clinical research and the interdependencies between these tasks' attributes.

The sample to test the hypotheses consists of 227 diabetes clinical trial observations between the years of 1981 and 2009. I have retrieved clinical trial data from the Federal Drug Administration's website that includes data on all clinical trials associated with treatments

seeking regulatory approval in the United States. Since the United States is the main market for pharmaceutical companies, the FDA database which is the most comprehensive of its sort covers a significant portion of the universe of all clinical studies carried across the globe. I have retrieved clinical study data by using relevant search words for identifying studies that are addressing Diabetes conditions and downloading all available information. After downloading the data, I have gone through 4569 clinical locations associated with the 227 trials listed in the dataset to correct for potential mistypes.

In my dataset I'm analyzing clinical trials sponsored by organizations located in the United States. Therefore, I'm referring to United States as the home country and international locations as those trial locations outside the United States. The clinical trial sites are dispersed over the globe; the United States has the most clinical sites located within its borders; 452348 out of 796003 sites listed in the FDA database are situated in the United States which corresponds to 56.8% but within my dataset, the ratio is slightly higher at 61%; 2787 sites out of 4569 are located in the United States followed by Canada, Germany, United Kingdom and France. These top 5 locations correspond to more than 80% of total clinical trial locations in diabetes trial subset.

5.3.2 Measures

Dependent Variable.

Timely Completion. To assess the impact of internationalization and coordination mechanisms on the timely completion of clinical trials, I have created a timely completion variable consisting of 8 categories in order of their distance in terms of standard deviation(s) from the mean completion time of all studies in the dataset. The studies that go into the first category are those which have a completion time that is more than 3 standard deviations above

the mean completion time (measured as number of days from the start date of the study to the completion date of the study) of all respective studies in the dataset. The categories follow in order of standard deviation difference, i.e. second category includes those studies that are concluded in a time that is below or equal to 3 standard deviations from the mean completion time, but above 2 standard deviations from the mean completion time. In total, this procedure generates 8 ordered categories; the latest completions with respect to the mean constituting the first category and the earliest completion with respect to the mean constituting the last (eighth) category.

In order to overcome the potential problem of mismatches in creating these categories, I have created respective categorical scales for the three phases that are observed in the dataset. For example, Phase I studies are "conducted to determine the metabolism and pharmacologic actions of drugs in humans, the side effects associated with increasing doses, and to gain early evidence of effectiveness" (FDA, 2010) and deal with initial efficacy and safety checks which take significantly less time to complete. Therefore each Phase I study's completion time is compared against the respective mean completion time and standard deviation distances of the population of Phase I studies rather than the measures for the entire set of clinical trials. The same procedure is applied for constructing the respective categories of Phase II and Phase III studies. Timely completion is an important measure of clinical trial success since sponsors of the clinical trial depend on FDA approval for marketing their drugs (Azoulay, 2004) and potential delays may cause losses in terms of monetary benefits and may even harm the reputation and market position of sponsoring firms.

Independent variables

Partner Complexity and Coordination. Using information from the content coding of sponsors collaborating in clinical trials, this variable refers to the total number of unique organizations/individuals that support the clinical trial. This variable captures the amount of coordination costs in terms of communication, information sharing and decision making efforts across sponsors of a given clinical trial. It is similar in nature to the measure used to capture cross-institutional coordination in distributed work (Cummings *et al.*, 2005, 2007).

Extent of Internationalization. I have calculated the number of unique countries where the clinical trial sites are located for a given clinical trial as a measure of the extent to which clinical trial project is internationalized.

Site Coordination Complexity. I have calculated the number of sites that are involved in a clinical trial study in order to measure the potential interdependence and coordination complexity associated with having multiple locations in a clinical study. This variable captures coordination problems that occur while managing sites across domestic and international site locations.

Same country sites (Collocation at country level). I have calculated the number of United States based clinical trial locations to capture the coordination benefits of having distributed project locations within the borders of the home country.

Control variables

Universities and Research Centers. Medical universities and the associated research centers conduct a significant part of their clinical trials in house and their studies are geared toward testing compounds from scientific discoveries for potential applications which creates a bias against clinical studies by these organizations with respect to the studies conducted by

pharmaceutical firms. To control for this effect, I introduce a dummy variable taking the value 1 if universities and research organizations are sponsoring the clinical trials and 0 otherwise.

Basic Science Based Studies. Studies designed to test the effect of an intervention based on discoveries of basic science in labs are involved with understanding the underlying mechanisms of the compounds and are not market oriented in nature. Hence, such studies may have a different time frame for application of the study. To control for the effect of such studies, I'm including a dummy variable that takes the value of 1 if the study design involves basic science and 0 otherwise.

In-house Sites. Sponsors of clinical trials may have capabilities to conduct the clinical trials in house in which case the underlying mechanism of coordination and control differs significantly. To control for this factor, I created a dummy variable taking the value of 1 if the clinical site is located in house and 0 otherwise.

NIH Organizations. The involvement of NIH related organizations such as "National Eye Institute" or "National Institute for Diabetes and Kidney Disorders" is a factor that may influence the study completion time because these organizations have unique resources such as a grand database of patients registered through government agencies and expertise in conducting clinical trials in house. Hence, to control for the effect of the involvement of such organizations, I have coded a dummy variable that takes the value of 1 if a government organization is sponsoring the study and 0 otherwise.

Treatment Based Studies. Studies designed to include new treatments which may involve new approaches to surgery or new combinations of drugs rather than existing treatment modes with new compounds are controlled for by a dummy variable that takes the value of 1 if the study is a treatment based study and 0 otherwise.

Randomized Studies. A randomized study is a study in which participants are randomly (i.e., by chance) assigned to one of two or more treatment arms of a clinical trial (FDA, 2010). Occasionally placebos are utilized. To control for potential effects brought by the modes of conduct between randomized trials and non-randomized trials, a dummy variable is introduced taking the value of 1 if the trial is randomized and 0 otherwise.

Cross-over Assignment Studies. A cross-over assignment refers to the alternating assignment of subjects across placebo and treatment groups over the duration of the study. To control for the effects of implementing this particular design on the timely completion of the given study, I include a dummy variable taking the value of 1 if the study includes a cross-over assignment design and 0 otherwise.

Enrollment. Clinical trial studies have different number of subjects involved across different phases. The involvement of a high number of subjects may increase potential problems in conducting the clinical trial, i.e., individuals not showing up, attrition problems etc. To control for the effect of this factor, I included the total number of patients enrolled in a clinical trial study.

5.3.3 Analysis

The dependent variable in my analysis is the timely completion of a clinical trial process as determined by the fit of the clinical trial completion time into one of the eight ordered categories that the study duration falls into. As such I am using ordered logistic regression to estimate the effects of my predictors. I also use OLS as a robustness check and control for the differences across phases, I receive similar results. Overall, my results are stable across both approaches with no significant changes in direction, magnitude or significance.

In Table 8, I report the descriptive statistics and correlations of the variables. Although the correlations reported are relatively low, I have derived the variance inflation factors (VIFs) as a control from the final model and the average VIFs were well below the generally accepted threshold of 10 (Greene, 1997). VIFs are typically made as a post estimation command for linear regressions, but as they make an assumption about the relationship between the independent variables, it is possible to use this technique for other functional forms as well (Menard, 2002). In testing my hypotheses, I also included the variables stepwise to check and make sure that the signs of coefficients are stable across the regressions. If multicollinearity would have been a major issue, sign and coefficients could have changed direction. Taken together, these precautions minimized the problem of multicollinearity.

TABLE 8: DESCRIPTIVE STATISTICS AND CORRELATIONS

	Variables	Mean	Std. Dev.	Min	Max	1	2	3	4	5	6	7	8	9	10	11	12	13
1	Timely Completion	-4.02	3.14	-8.00	-1.00	1												
2	Universities and Research Centres	0.30	0.46	0.00	1.00	-.36	1											
3	Basic Science Based Studies	0.01	0.10	0.00	1.00	-.10	.03	1										
4	In house sites	0.31	0.46	0.00	1.00	-.16	.29	.03	1									
5	NIH funded studies	0.23	0.42	0.00	1.00	-.14	-.12	-.06	.09	1								
6	Treatment Based Studies	0.85	0.36	0.00	1.00	.26	-.29	-.25	-.05	.05	1							
7	Randomized Studies	0.91	0.28	0.00	1.00	.08	.09	.04	-.33	-.43	-.11	1						
8	Cross-over Assignment	0.11	0.31	0.00	1.00	.27	.02	.14	-.04	-.14	.10	.12	1					
9	Enrollment	875.10	4297.1	0.00	61201	.08	-.06	-.01	.10	.14	.03	-.22	-.03	1				
10	Partner Complexity	1.55	1.10	1.00	7.00	-.37	.52	-.06	.12	.05	-.20	.08	-.08	-.05	1			
11	Extent of Internationalization	1.96	3.58	1.00	38.00	-.11	-.18	.05	-.16	-.14	-.09	.12	-.11	-.03	.00	1		
12	Site Coordination Complexity	11.69	42.24	1.00	476.00	-.08	-.18	-.01	-.02	-.16	.03	.11	-.10	-.02	-.06	.46	1	
13	Same country sites (Collocation)	6.30	19.81	0.00	171.00	.13	-.28	-.04	.03	-.15	.15	.11	-.10	-.02	-.12	.05	.44	1

TABLE 9: ORDERED LOGISTIC REGRESSION MODELS OF TIMELY COMPLETION

Variable	Model 1 Timely Completion	Model 2 Timely Completion	Model 3 Timely Completion	Model 4 Timely Completion	Model 5 Timely Completion
Universities and Research Centres	-1.595***	-1.119***	-1.416***	-1.577***	-1.503***
	[0.372]	[0.421]	[0.439]	[0.462]	[0.471]
Basic Science Based Studies	-1.844	-2.086	-1.895	-1.974	-1.822
	[1.367]	[1.406]	[1.429]	[1.418]	[1.476]
In house sites	-0.0406	-0.0531	-0.159	-0.0752	-0.081
	[0.363]	[0.370]	[0.374]	[0.378]	[0.380]
NIH funded studies	-0.947**	-0.893**	-1.158***	-1.281***	-1.293***
	[0.400]	[0.411]	[0.432]	[0.448]	[0.451]
Treatment Based Studies	0.637	0.648	0.587	0.613	0.543
	[0.459]	[0.466]	[0.469]	[0.471]	[0.473]
Randomized Studies	0.264	0.313	0.351	0.435	0.404
	[0.538]	[0.553]	[0.558]	[0.565]	[0.565]
Cross-over Assignment	2.306***	2.230***	2.105***	2.063***	2.065***
	[0.715]	[0.711]	[0.719]	[0.725]	[0.720]
Enrollment	8.17E-05	8.14E-05	8.32E-05	8.31E-05	8.28E-05
	[3.61e-04]	[3.86e-04]	[4.09e-04]	[4.10e-04]	[3.98e-04]
Partner Complexity		-0.374**	-0.340**	-0.330**	-0.330**
		[0.161]	[0.161]	[0.160]	[0.162]
Extent of Internationalization			-0.0979**	-0.0718*	-0.0157
			[0.0390]	[0.0400]	[0.0494]
Site Coordination Complexity				-0.00938	-0.0470*
				[0.00690]	[0.0242]
Same country sites (Collocation)					0.0489*
					[0.0279]
Log Likelihood	-224.1	-221.1	-217.5	-216.1	-214.2

* p<.1, ** p<.05 ***p<.01; two-tailed tests. Standard errors are in brackets. 227 Clinical trial observations.

5.4 Results

Table 9 shows the results of the ordered logistic regression analysis where the dependent variable is the timely completion of clinical trials. Model 1 shows the baseline model which includes controls for involvement of NIH related organizations in the study, inclusion of universities and research centers, number of patients enrolled, in house trial locations, study designs involving basic science, new treatments, randomized trials and cross-over assignments. In Model 2, I introduce the first of my theory variables of interest (partner complexity). Model 3 includes the second theory variable (extent of internationalization), Model 4 includes (site complexity and coordination) and Model 5 includes collocation variable (same country based sites). Each model represents a significant improvement over the baseline model and the prior model with the log likelihood value increasing from –224.1 for the base model ($p<.001$) to –214.2 ($p<.001$) for the final model represented in Model 5 of Table 9. The baseline model was generally consistent with prior research findings for clinical trials. In line with previous research findings, the basic science based studies have a negative effect on the timely completion variable (Huckman *et al.*, 2008).

Hypothesis 1: *Partner Complexity.* I argued that the partner complexity and the associated costs of coordination as captured by the number of sponsors on a given trial will negatively effect the timely completion of the clinical trial. In Model 2, the coefficient for partner complexity variable was statistically significant. The main effect of the task breadth variable was negative ($b=-0.374$, $p<.05$). This result is significant throughout the rest of the models and supports my hypothesis: there are costs to having more sponsors involved in clinical trial process which reflect itself negatively on the timely completion of the clinical trial. Moreover, analyzing the change in probabilities for timely completion while keeping other

variables at their mean, a one standard deviation change in partner complexity increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 1.9% while it decreases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 4.4%.

Hypothesis 2: *Extent of Internationalization.* I predicted that the extent of internationalization of distributed work projects as represented by the number of unique countries involved in a clinical study would have a negative effect on the timely completion of the clinical trial. Given Model 3's improvement over prior model, I inspect the coefficient of extent of internationalization variable which is significant and negative (the coefficient is significant; $b=-0.0979$, $p<.05$). The result holds consistently in Model 3 but this effect disappears in further models. This variable points to the negative effect of coordination of internationally distributed work project on timely completion, giving partial support for Hypothesis 2. Moreover, analyzing the change in probabilities for timely completion while keeping other variables at their mean, a one standard deviation change in extent of internationalization increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 10.1% while it decreases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 5.9%.

Hypothesis 3: *Site Coordination Complexity.* I argued that the coordination complexity due to interdependencies across sites as measured by the number of clinical trial sites will have a negative effect on the timely completion of the study. Given Model 4's improvement over prior models, I inspect the coefficient of site coordination complexity variable which is negative as

expected yet not significant. The coefficient becomes significant and negative in Model 5 (b=-0.047, p<.10). Given this result, I find partial support for Hypothesis 3, there is a negative relationship between number of sites involved in a clinical study and the timely completion of the study. Analyzing the change in probabilities for timely completion while keeping other variables at their mean, a one standard deviation change in number of sites involved in a clinical study increases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 73.8% while it increases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 55.4%.

Hypothesis 4: *Same country sites (collocation at country level).* I predicted that having distributed work project locations at home country would allow for reduction in coordination costs and increase efficiency of distributed work. In Model 5, the coefficient for task same country variable was statistically significant and positive as expected (b=0.0671, p<.05). This result supports my hypothesis: there are benefits to having sites in home country; the effect of which is to reduce coordination costs and increase efficiency in distributed work. Looking at the changes in probabilities for timely completion, it is possible to see that while keeping other variables at their mean, a one standard deviation change in same country variable decreases the likelihood of being in the first category (being among the latest to complete the trial with respect to other trials in the same phase) by 10.1% while it increases the likelihood of being in the last category (i.e., being among the first to complete the trial with respect to the other trials in the same phase) by 26.4%.

5.5 Discussion

The role of coordination in internationalization of the research and development tasks in distributed work processes is understudied, even though recent studies systematically document that geographical dispersion, team familiarity and task complexity effect the performance of distributed work projects such as software development (Espinosa *et al.*, 2007) as well as patent examination process (Harhoff *et al.*, 2009) and contractual alliances (White *et al.*, 2005). The primary issue of concern in this study is to disentangle the effects of internationalization, collocation benefits accruing from being in the same country, partner complexity and coordination complexity on the outcome of geographically distributed work processes. This study documents how different attributes of internationalization of the distributed work create complexities which can hinder the performance of a distributed work project and cause potential losses in terms of time and money. Below, I explain the implications of these findings.

First, I find evidence for negative effects of partner complexity on project outcomes in terms of the timely completion of projects. My findings build up on previous literature and further tease out the effects of coordination problems that hamper collaborations in joint work efforts (Cummings *et al.*, 2005, 2007; Gulati *et al.*, 1998; Kiesler *et al.*, 2002). Attending to such problems is important as the number of collaborations in sponsoring clinical trials has increased steadily over the last decade, driven by several different motives including the need to collaborate and innovate in a rapidly developing field (Powell *et al.*, 1996; Rothaermel *et al.*, 2006; Rothaermel *et al.*, 2007). Given the results of this study, the partner complexity and associated coordination costs entail an issue of concern for organizations willing to jointly sponsor clinical trials.

Second, I supply empirical evidence for the effects of internationalization of distributed work and its attendant coordination problems on project performance in clinical drug trials in diabetes. My primary contribution lies within the emphasis of management of joint work projects across borders and how coordination of distributed work across countries impacts project outcomes. By analyzing the impact of temporal, spatial and cultural distance on the coordination and timely completion of distributed work, I bring to the fore a more fine-grained understanding of factors influencing the effective coordination of internationally distributed research and development projects.

And third, I was able to show the effect collocation as a coordination mechanism in mitigating the problems associated with international distribution of work projects. Having a higher number of locally based units involved in distributed work projects reduces the efforts for coordination and enhances the performance of distributed work projects. This is an important finding given that the collocation can still be an important coordination mechanism in a setting where coordination by planning in design is found to be very effective.

5.5.1 Limitations

In spite of the study's contributions, limitations exist. I follow a reduced form approach to investigate the effects of internationalization, partner complexity and collocation on the actual outcomes rather measuring directly intermediate processes and mechanisms. I could not have access to intermediate level data to actually measure the duration of task routines and problems that arise during clinical trial processes. Also, I do not have a control for the clinical site activity outside the diabetes research which may be influencing the outcomes of clinical trials in the dataset. Inclusion of further data on other trials would help solve this problem.

5.6 Conclusion

Taken as a whole, my results provide evidence that different attributes of internationally distributed work in terms of partner complexity, internationalization and collocation effect the coordination and timely completion of clinical trial projects. Internationalization and coordinative complexity of tasks increase the time to completion where as collocation as a coordination by feedback mechanism helps with the effective management of distributed work projects and improves project performance. At the same time, having multiple sponsors increases coordination problems that originate from management of communication and joint problem solving efforts and hampers the performance of distributed projects. More importantly, a deeper understanding of the effect of collocation within the country of sponsor has been analyzed. Collocation increases the ability to communicate more effectively and has been shown to be positively associated with the performance of distributed work projects, but the extent to which collocation can be beneficial above and beyond coordination by planning had not been documented. Future research could look further into details and the nature of communication and task assignment as they relate to distribution of the projects and how collocation can help mitigate concerns. Moreover, I chose country level collocation as a measure to study the effectiveness of internationalization of trials; future research could look at other measures of collocation such as geographic areas, cities, zip codes etc., and see if there are reasonable changes in effect sizes by use of smaller distance units or if country level is in fact the most prominent level at which collocation mechanism can be seen.

Overall, this study contributes to theories of management in international R&D, coordination in distributed work and knowledge sharing across borders. Although previous research has attended to issues of coordination in international R&D, this research makes a

unique contribution by looking at the management of distributed work and how internationalization effects the coordination of such work above and beyond the already accepted effects of geographical dispersion on coordination. More precisely, it shows the additive effect of coordination by feedback mechanism that is enhanced through collocation and how this coordination mechanism can help overcome problems stemming from distribution of work.

CHAPTER 6

Conclusion

In this dissertation I address the central question of how learning from prior experience mitigates coordination concerns in collaborative projects. In doing so, I explicitly address the differential effects of learning from various types of joint experience and how these different types of experience effect coordination of activity. Prior research has analyzed resolution of coordination problems in collaborative work from the perspectives of learning, capability development and design. However, scant attention has been paid to the intersection of the two; namely how design of prior collaborations effect the build up of an ability to coordinate action through joint work. By specifically focusing on this junction, I analyze the differential effects of learning from collaborations given project characteristics and the impact of such learning on future projects. For this dissertation I have managed to find two settings in which I can observe elements of collaboration in joint work both within teams and across organizations.

In Chapter 2, I use patent and publication data in addition to interviews with scientists in the field to study how collaborations differ within and outside a technology domain where a major disruptive change occurs, such as the emergence of research in hESC area. I use interviews to get insights from scientists working in stem cell labs to gain a better understanding of the field and how collaborations are managed in this research area. In addition, I use patent and publication data as my primary data sources. My data includes all work in this domain and is not limited to the US, as was the case in most previous studies focusing on scientific collaborations. This ensures that I have a complete picture of the relationships in this domain. I find that prior experience in collaborative work in a setting that changes rapidly may actually be detrimental to performance of the focal project.

In Chapter 3, I use data from clinical trials in the diabetes disease area and analyze the design of distributed work in clinical trials. I am able to observe the entire set of clinical trials reported to the FDA. I find that task complexity, breadth and partner complexity increase coordination problems and negatively affect the performance of projects.

In Chapter 4, I use data from clinical trials in five different disease areas, supplemented with archival data on clinical research organizations and trial sites. I particularly construct variables on collaborative experience in overall work as well as collaborations specific to the disease area and to the interventions involved. I find that the effect of collaborative experience in solving coordination problems in a distributed work setting is enhanced when experience in joint work is based on the same knowledge domain as the focal project. Moreover, I find that working on complex projects enhances the ability of organizations' to coordinate action.

In Chapter 5, I use clinical trials in diabetes area supplemented with information on internationalization of clinical trial locations to test my hypotheses. I particularly look at coordination strategies using collocation as a way to solve potential coordination problems. I find that even in distributed work projects that focus on well a codified knowledge area such as that of diabetes where protocols of the trial should be more easily executed and where modularization, coordination by planning, should be easier to achieve, there are still substantial benefits to collocation as can be observed from the effect of collocation on project performance.

6.1 Main Findings

In this section I observe the empirical findings across chapters and provide answers to the main research questions of this dissertation.

6.1.1 The Role of Learning from Experience in Scientific Teams

In Chapter 2, I documented how the benefits of collaborative experience make a difference on the scientific impact of the discovery itself. Further, I find evidence that prior experience has diverse effects on mitigating coordination costs stemming from scientific distance (knowledge distance across individual inventors on a patent team) and departmental distance. I supply empirical evidence for the effects of collaborative experience and its attendant coordination problems on scientific impact. My primary contribution lies within the emphasis of coordination costs that constitute an important barrier on the effectiveness and efficiency of joint work in innovative settings. A plethora of research has documented the benefits that accrue from collaborating with partners (Cummings *et al.*, 2005, 2007; Fleming *et al.*, 2007; Guimera *et al.*, 2005; Jones *et al.*, 2008; Powell *et al.*, 1999; Powell *et al.*, 1996). A much less attended issue that I bring to the fore is the coordination costs that emerge when there are geographical, institutional and cultural barriers to bridge that require ongoing dialogue and negotiation (Cummings *et al.*, 2005, 2007, 2008; Grant, 1996). By combining the collaborative benefits and coordination costs, one can reach a more fine-grained understanding the conditions that allow for successful collaborations.

Second, I was able to analyze the effects of coordination costs on the research outcomes of collaborative work in human embryonic stem cells. Having inventors from different departments on a patent team increases the costs of coordination as it increases the complexity and difficulty of communication and task allocation in joint work. It is harder for teams with inventors from multiple departments to bridge institutional gaps and have a shared social setting or maintain awareness of what others are doing. However, in the existence of prior work

relationships, coordination problems are overcome and departmental diversity enhances knowledge creation efforts.

Third, heterogeneity among the inventors on a patent team was expected to have a positive and diminishing effect on the impact of the resulting invention because the ability of the team to recombine knowledge peaks when the there is enough absorptive capacity across the team so that members can understand each other's work while being distant enough to bring in new perspectives to the solution domain. (Owen-Smith *et al.*, 2003). The results show that the scientific distance to impact relationship among the inventors of a patent does show an inverted-U type curvilinear shape as expected. However, when team members have prior joint work experience outside of hESC domain, the effect of distance is a U-shaped curve where either low or high levels of distance are optimal. This may be because in the existence of prior experience outside hESC, teams build work routines that aren't applicable to hESC domain; the existence of such routines creates rigidities which moderate the relationship between scientific distance and knowledge creation in hESC.

Finally, I find evidence for positive effects of prior experiences in patenting in a given area, hESC in this case, are hampered when the activities of the research team have to be coordinated across multiple departments and across geographic space. The learning literature has made great efforts to investigate how individuals and organizations draw inferences from past experiences (Levitt *et al.*, 1988). My findings suggest that experiential learning is greatest when the team is co-located. Compared to interdepartmental teams, local teams can arguably spend more time discussing and resolve potential conflicts that arise. By working in the same department, tacit knowledge may spill over to other team members enhancing the possibility for the team to improve performance.

6.1.2 Learning and Coordination in Distributed Work

In Chapters 3 and 4, I analyze the effects of design characteristics on project performance and the role of learning from experience in mitigating coordination problems in inter-organizational relations, particularly in distributed work settings. In Chapter 3, I focus on the design of distributed work projects and show that complexities stemming from task size, task breadth and involvement of multiple partners are detrimental to effective coordination and management of distributed work projects and can hamper performance. In Chapter 4, I look at the link between learning that takes place at the inter-organizational level and how different design characteristics pertaining to the knowledge domain of the project and the complexity of the project affect the build up of coordination ability. My findings give insight into both the design and the management of distributed work projects

First, I supply empirical evidence for the effects of design characteristics in terms of task size, task breadth, task applicability and the involvement of multiple partners in defining the complexity of the project. I find that as complexity stemming from task size, task breadth and project dispersion increases, the associated coordination costs surge. My primary contribution lies within the emphasis of coordination costs that constitute an important barrier to the effectiveness and efficiency of distributed work.

Second, I show the positive effects of learning from prior experiences in distributed work projects and how such learning helps overcome potential coordination problems in focal projects. Especially the prior joint work between lead sponsor and locations (clinical trial sites) has an important positive effect on the success of focal projects. This effect is valid not just for the joint

work overall, but exists strongly in joint work building on similar knowledge and joint work in complex projects as well.

Third, I find evidence for the role of knowledge relevance and project complexity in enhancing the importance of prior experiences as such experiences help build a coordination ability to resolve potential problems in focal projects. This finding is relevant when one thinks of the layers of experience that are more pronounced in relation to the knowledge domain and task complexity. Although, overall experience in joint work has been shown to be an important factor in mitigating coordination problems, a more fine-grained look at the effect of experience across levels shows that, when the experience in joint work is based on the same knowledge domain as that of the focal project, the impact of this experience in mitigating coordination problems is higher than experience in joint work from other knowledge domains. At the same time, when joint work involves projects with complex tasks, work in such projects helps develop a greater ability to coordinate action among distributed work project partners than in cases where the joint work projects don't involve such levels of complexity.

6.1.3 Coordination Mechanisms and Collocation in Distributed Work

In Chapter 5, I analyze the coordination mechanisms in the organization of distributed work and I particularly focus on collocation as a means to resolve coordination problems. Although recent studies point to the effects of geographical dispersion, team familiarity and task complexity on the performance of distributed work in settings such as software development (Espinosa *et al.*, 2007), patent examination process (Harhoff *et al.*, 2009) and contractual alliances (White *et al.*, 2005), the particular role of coordination mechanisms in mitigating coordination concerns is understudied. In this study, I find evidence for the positive effect of collocation in a distributed work setting where coordination by plan would be the expected norm.

I show that there are benefits to collocating in resolving coordination problems beyond what can be offered through the coordination by plan mechanism.

I also supply empirical evidence for the effects of internationalization of distributed work and its attendant coordination problems on project performance in distributed work projects. My primary contribution lies on the emphasis of location choices in distributed work projects, design choices that determine project attributes and the associated coordination mechanisms that impact project outcomes. I find that internationalization of trials has a negative effect on performance while complexities stemming from having to coordinate action across multiple partners also hinder project performance. On the other hand, I find evidence for the positive role of collocation as a coordination mechanism in helping overcome coordination problems. By analyzing the impact of temporal, spatial and cultural distance on the coordination and timely completion of distributed work, I bring to the fore a more fine-grained understanding of factors influencing the effective coordination of distributed work projects from a coordination and internationalization perspective.

6.2 Implications for Theory

6.2.1 The Role of Learning from Experience in Scientific Teams

In Chapter 2, I build up on previous literature and further tease out the effects of coordination problems that hamper interdepartmental collaborations (Cummings *et al.*, 2007). Attending to such problems is important as the number of interdepartmental teams has increased steadily over the last 25 years (Jones *et al.*, 2008). This trend is driven by several different motives including access to new instrumentation (de Solla Price, 1986), reduced communication costs enabled through new communication technologies (Agrawal *et al.*, 2008), complementarities of knowledge and experience to generate new scientific insights (Basalla,

1988) and solving increasingly complex problems that would be intractable for a single individual to solve. Taken as a whole, my results proffer evidence that team level joint production experience and team structure in terms of scientific and departmental distance are important determinants of patent impact. Having scientists from multiple departments reduces the impact of the patented invention; however, having medium level of scientific distance across inventors enhances the positive effect of recombination on patent impact. At the same time, learning from experience appears to be an important factor determining the patent outcomes. More importantly, although much enthusiasm surrounds interdepartmental collaborations, an issue that has been somewhat neglected was the coordination costs that constitute the downside of collaboration.

In this chapter, I am able to show that not all types of experience can be useful in mitigating coordination concerns. Especially in settings where radical changes occur, prior experience in joint and the routines learnt in coordinating action may not be beneficial. As such this chapter raises some theoretical concerns regarding the organization of teams where knowledge is dispersed across members and where a disruptive change in the field may render not just the prior knowledge but the prior experiences in joint work obsolete.

6.2.2 Learning and Coordination in Distributed Work

In Chapter 3, I build on the literature in distributed work and show that task characteristics such as complexity, size and breadth impact the coordination and effective management of distributed work projects. Further, analyzing the role of experience in inter-organizational settings is a well studied topic; however there are important caveats to studying experience as an overall measure rather than analyzing the different levels of experience and the underlying logic with which experience impacts performance of collaborative work. The primary

issue is one of harnessing the benefits of experience (by knowledge relevance and task complexity) while measuring the impact of potential coordination problems (by including measures of sponsor, location and protocol specific complexities). In Chapter 4, I document how the benefits of different types of collaborative experience impact the effective management of distributed work projects. Further, I find evidence that prior experience in joint work not only positively effects the project performance but also the effectiveness of these experiences is enhanced when the ties' experience stems from similar knowledge domains as that of the current project, perhaps because more relevant tasks are involved in the current project or because learning based on specialization is especially important in this complex setting.

6.2.3 Coordination Mechanisms and Collocation in Distributed Work

In Chapter 5, I focus on the coordination mechanisms that are used distributed work projects. My findings build up on previous literature and further tease out the effects of coordination problems that hamper collaborations in joint work efforts (Cummings *et al.*, 2005, 2007; Gulati *et al.*, 1998; Kiesler *et al.*, 2002). Attending to such problems is important as the number of collaborations in sponsoring has increased steadily over the last decade, driven by several different motives including the need to collaborate and innovate in rapidly developing fields (Powell *et al.*, 1999; Powell *et al.*, 1996; Rothaermel *et al.*, 2006; Rothaermel *et al.*, 2007).

I show that in addition to the mechanism of coordination through plan, the coordination through feedback also has a positive effect on the project performance even in the existence of a planning mechanism that allows for an efficient distribution and control of work up-front.

6.3 Implications for Practice

The findings from this thesis point to some important areas where policy makers and practitioners can draw insights from. First of all, my study in team work and coordination in a

dynamic setting shows that although moderate scientific distance in knowledge based teams may be beneficial, one has to be careful of the prior experiences of team members. If members of the team draw conclusions regarding routines and work processes to employ, based their prior experiences that are not applicable to the current setting, performance of the team could be hampered significantly.

Second, I analyze the design of complex projects in distributed work environments by looking at the various characteristics of tasks involved in clinical trials. In Chapter 3, I show that the inclusion of multiple tasks that vary in size and breadth could be highly taxing for project coordination and hamper performance. Although I don't have data on project costs to make a clear-cut comparison between project costs and task inclusion, one can assume that bundling of multiple tasks in distributed work projects brings about complications with regards to the efficiency with which distributed work can be conducted.

Third, I show that there are significant benefits to selecting partners who have not just worked in the same setting but also worked in projects involving a similar knowledge base. Moreover, the benefits of learning from such joint work seem to increase significantly in project complexity. Both of these factors should alert practitioners to the importance of prior experience in partner selection where other considerations regarding the cost of the project may blur the picture. Delays in project completion may prove more harmful than the costs of contracting with clinical trial sites.

Finally, I show evidence that collocation still holds as an important mechanism in a setting where planning and modularization can be expected due to the codified nature of the knowledge. This study shows that there are benefits to collocating beyond what can be

anticipated in the design of distributed work projects. This finding has implications for not only the design of clinical trials but also for the selection of the location of clinical trial sites.

6.4 Future Research Extensions

Chapter 2 deals with organization of scientific teams and the discovery process, taking a knowledge recombination and coordination perspective. The findings suggest that diversity may be a necessary factor for knowledge generation but managing diversity can be difficult especially in cases where the ability to coordinate action across partners with diverse knowledge is limited. Although the findings of this research are valuable for the organization of scientific teams, there are also several areas on which future research could improve. First, future research could analyze how the scientific distance between members of scientific teams plays out in contexts where there is no fundamental change in the underlying knowledge related to the focal area. Analyzing prior experience in joint work and how the prior experience affects resolution of coordination problems would give a more profound understanding of the link between optimum structures of scientific teams in changing environments. Second, further data in intermediate steps of coordination on the day to day level could help shed more light on the process of scientific discovery as it relates to team structure and coordination problems. Third, the experience measure that I use varies in terms of the knowledge domain at which the experience occurs, through a controlled experiment it may be possible to vary experience and ability to coordinate differentially.

My findings from Chapter 3 and 4 relate to the design of distributed work in terms of task characteristics and how experience with projects that have certain design complexities help mitigate coordination concerns in further projects. My findings in Chapter 3 relate to how design of the distributed project effects coordination and performance of the project. In Chapter 4, I

build up on this area and show that experience in complex projects and in projects that build up on similar knowledge bases has a positive effect on performance which is attributable to the fact that such experience helps resolve coordination problems. A further step in this line of research would be to look at how distribution of work affects the build up of a coordination capability by looking at intermediate processes and build up of routines. Further, one can analyze the extent to which partners learn from each other in coordinating distributed work projects, i.e., how does experience with outside parties effect coordination of action with focal partner? Moreover, analyzing data on trial costs can give a better picture of how delays originating from choices in project design affect the performance of sponsoring organizations.

In Chapter 5, I particularly look at the design of distributed work projects from a coordination perspective and tease out the effects of collocation. Ideally, collocation as a coordination mechanism (coordination by feedback), should be useful as it allows for effective communication and timely adjustments to performance. I find that even in a well codified knowledge setting (Type II Diabetes Trials), coordination by feedback is still a valuable mechanism above and beyond what coordination by plan mechanism can offer. Going further, researchers can analyze to what extent coordination by feedback is useful in settings where planning is possible through a comparative study across areas where knowledge codification and the ability to effectively modularize projects varies.

References

Agarwal R, Ganco M, Ziedonis RH. 2009. Reputations for toughness in patent enforcement: Implications for knowledge spillovers via inventor mobility. *Strategic Management Journal* **30**(13): 1349-1374

Agrawal A, Goldfarb A. 2008. Restructuring research: communication costs and the democratization of university innovation. *The American Economic Review* **98**(4): 1578-1590

Ahuja G. 2000. Collaboration Networks, Structural Holes, and Innovation: A Longitudinal Study. *Administrative Science Quarterly* **45**(3): 425-455

Ahuja G, Morris Lampert C. 2001. Entrepreneurship in the large corporation: A longitudinal study of how established firms create breakthrough inventions. *Strategic Management Journal* **22**(6 7): 521-543

Aiken LS, West SG. 1991. *Multiple regression: Testing and interpreting interactions*. Sage Publications Inc

Albert MB, Avery D, Narin F, McAllister P. 1991. Direct validation of citation counts as indicators of industrially important patents. *Research Policy* **20**(3): 251-259

Anand BN, Khanna T. 2000. Do firms learn to create value? The case of alliances. *Strategic Management Journal*(21): 295-315

Arrow KJ. 1962. The economic implications of learning by doing. *The Review of Economic Studies*: 155-173

Autio E, Kanninen S, Gustafsson R. 2008. First-and second-order additionality and learning outcomes in collaborative R&D programs. *Research Policy* **37**(1): 59-76

Autio E, Sapienza HJ, Almeida JG. 2000. Effects of age at entry, knowledge intensity, and imitability on international growth. *The Academy of Management Journal* **43**(5): 909-924

Azoulay P. 2004. Capturing Knowledge within and across Firm Boundaries: Evidence from Clinical Development. *The American Economic Review* **94**(5): 1591-1612

Azoulay P, Stellman A, Zivin JG. 2006. PublicationHarvester: An open-source software tool for science policy research. *Research Policy* **35**(7): 970-974

Banker RD, Davis GB, Slaughter SA. 1998. Software development practices, software complexity, and software maintenance performance: A field study. *Management Science*: 433-450

Barkema HG, Shenkar O, Vermeulen F. 1997. Working abroad, working with others: How firms learn to operate international joint ventures. *The Academy of Management Journal* **40**(2): 426-442

Basalla G. 1988. *The evolution of technology*. Cambridge Univ Pr

Bechky BA. 2006. Gaffers, gofers, and grips: Role-based coordination in temporary organizations. *Organization Science* **17**(1): 3-21

Bensaou M, Venkatraman N. 1995. Configurations of interorganizational relationships: a comparison between US and Japanese automakers. *Management Science*: 1471-1492

Bercovitz J, Feldman M. 2011. The mechanisms of collaboration in inventive teams: Composition, social

networks, and geography. *Research Policy* **40**: 81-93

Bresman H, Birkinshaw J, Nobel R. 1999. Knowledge transfer in international acquisitions. *Journal of International Business Studies* **30**(3): 439-462

Burke K, Aytes K, Chidambaram L, Johnson JJ. 1999. A study of partially distributed work groups: The impact of media, location, and time on perceptions and performance. *Small Group Research* **30**(4): 453

Camerer C, Russell Sage F. 2003. *Behavioral game theory: Experiments in strategic interaction*. Princeton University Press Princeton, NJ

Cameron AC, Trivedi PK. 1986. Econometric models based on count data: comparisons and applications of some estimators and tests. *Journal of applied econometrics*: 29-53

Campbell DJ. 1988. Task complexity: A review and analysis. *The Academy of Management Review* **13**(1): 40-52

Clark HH. 1996. Using language. *Computational Linguistics* **23**(4)

Clark HH, Brennan SE. 1991. Grounding in communication.

Cramton CD. 2001. The mutual knowledge problem and its consequences for dispersed collaboration. *Organization Science* **12**(3): 346-371

Cummings JN, Kiesler S. 2005. Collaborative Research Across Disciplinary and Organizational Boundaries. *Social Studies of Science* **35**(5): 703

Cummings JN, Kiesler S. 2007. Coordination costs and project outcomes in multi-university collaborations. *Research Policy* **36**(10): 1620-1634

Cummings JN, Kiesler S. 2008. Who collaborates successfully?: prior experience reduces collaboration barriers in distributed interdisciplinary research: 437-446. ACM New York, NY, USA

de Solla Price DJ. 1986. *Little science, big science--and beyond*. Columbia University Press New York

Dennis AR, George JF, Jessup LM, Nunamaker Jr JF, Vogel DR. 1988. Information technology to support electronic meetings. *MIS Quarterly* **12**(4): 591-624

Dhanaraj C, Lyles MA, Steensma HK, Tihanyi L. 2004. Managing tacit and explicit knowledge transfer in IJVs: the role of relational embeddedness and the impact on performance. *Journal of International Business Studies* **35**(5): 428-442

Doz YL. 1996. The evolution of cooperation in strategic alliances: initial conditions or learning processes? *Strategic management journal* **17**(S1): 55-83

Dyer JH, Singh H. 1998. The relational view: Cooperative strategy and sources of interorganizational competitive advantage. *Academy of management review*: 660-679

Edmondson AC, Bohmer RM, Pisano GP. 2001. Disrupted routines: Team learning and new technology implementation in hospitals. *Administrative Science Quarterly* **46**(4): 685-716

Espinosa JA, Slaughter SA, Kraut RE, Herbsleb JD. 2007. Familiarity, complexity, and team performance in geographically distributed software development. *Organization Science* **18**(4): 613-630

Ethiraj S, Puranam P. 2004. The Distribution of R&D Effot in Systemic Industries: Implications for Competitive Advantage. *Advances in Strategic Management* **21**: 225-253

Faraj S, Xiao Y. 2006. Coordination in Fast-Response Organizations. *Management Science* **52**(8): 1155

FDA. 2010. Clinical Trial Information, Vol. 2010:

Feinberg SE, Gupta AK. 2004. Knowledge spillovers and the assignment of R&D responsibilities to foreign subsidiaries. *Strategic Management Journal* **25**(8-9): 823-845

Fleming L. 2001. Recombinant uncertainty in technological search. *Management Science* **47**(1): 117-132

Fleming L. 2002. Finding the organizational sources of technological breakthroughs: the story of Hewlett-Packard's thermal ink-jet. *Industrial and Corporate Change* **11**(5): 1059-1084

Fleming L, Mingo S, Chen D. 2007. Collaborative Brokerage, Generative Creativity, and Creative Success. *Administrative Science Quarterly* **52**(3): 443-475

Fleming L, Sorenson O. 2004. Science as a map in technological search. *Strategic Management Journal* **25**(8-9): 909-928

Galbraith JR. 1977. *Organization design*. Addison-Wesley

Ganco M. 2008. The influence of technological interdependence on employee entrepreneurship and mobility: Evidence from the semiconductor industry. *Atlanta*

Gassmann O, von Zedtwitz M. 1999. New concepts and trends in international R&D organization. *Research Policy* **28**(2-3): 231-250

George G, Kotha R, Zheng YF. 2008. Entry into Insular Domains: A Longitudinal Study of Knowledge Structuration and Innovation in Biotechnology Firms. *Journal of Management Studies* **45**(8): 1448-1474

Gerwin D. 2004. Coordinating new product development in strategic alliances. *The Academy of Management Review* **29**(2): 241-257

Gibson C, Vermeulen F. 2003. A healthy divide: Subgroups as a stimulus for team learning behavior. *Administrative Science Quarterly* **48**(2): 202-239

Gittelman M, Kogut B. 2003. Does good science lead to valuable knowledge? Biotechnology firms and the evolutionary logic of citation patterns. *Management Science* **49**(4): 366-382

Grant RM. 1996. Toward a Knowledge-Based Theory of the Firm. *Strategic Management Journal* **17**: 109-122

Greene WH. 1997. *Econometric analysis*. Prentice Hall Upper Saddle River, NJ

Guimera R, Uzzi B, Spiro J, Amaral LAN. 2005. Team Assembly Mechanisms Determine Collaboration Network Structure and Team Performance. *Science* **308**(5722): 697-702

Gulati R, Lavie D, Singh H. 2009. The nature of partnering experience and the gains from alliances. *Strategic Management Journal* **30**(11): 1213-1233

Gulati R, Singh H. 1998. The Architecture of Cooperation: Managing Coordination Costs and Appropriation Concerns in Strategic Alliances. *Administrative Science Quarterly* **43**(4): 781-784

Gulati R, Westphal JD. 1999. Cooperative or controlling? The effects of CEO-board relations and the content of interlocks on the formation of joint ventures. *Administrative science quarterly* **44**(3): 473-506

Gupta AK, Wilemon D. 1996. Changing Patterns in Industrial R&D Management. *Journal of Product Innovation Management* **13**(6): 497-511

Hall BH, Jaffe A, Trajtenberg M. 2005. Market value and patent citations. *Rand Journal of Economics*: 16-38

Hargadon A. 2003. *How Breakthroughs Happen: The Surprising Truth about How Companies Innovate*. Harvard Business School Press

Hargadon A, Sutton RI. 1997. Technology Brokering and Innovation in a Product Development Firm. *Administrative Science Quarterly* **42**(4): 716-749

Harhoff D, Wagner S. 2009. The Duration of Patent Examination at the European Patent Office. *Management Science* **55**(12): 1969-1984

Heath C, Staudenmayer N. 2000. Coordination neglect: How lay theories of organizing complicate coordination in organizations. *Research in Organizational Behavior* **22**: 153-192

Henderson R, Cockburn I. 1994. Measuring competence? Exploring firm effects in pharmaceutical research. *Strategic management journal* **15**(S1): 63-84

Hennart JF, Zeng M. 2005. Structural determinants of joint venture performance. *European Management Review* **2**(2): 105-115

Hinds P, Kiesler S. 2002. *Distributed work*. The MIT Press

Hinds PJ, Mortensen M. 2005. Understanding conflict in geographically distributed teams: The moderating effects of shared identity, shared context, and spontaneous communication. *Organization Science* **16**(3): 290-307

Hitt MA, Dacin MT, Levitas E, Arregle JL, Borza A. 2000. Partner selection in emerging and developed market contexts: Resource-based and organizational learning perspectives. *The Academy of Management Journal* **43**(3): 449-467

Hoang H, Rothaermel FT. Leveraging internal and external experience: exploration, exploitation, and R&D project performance. *Strategic Management Journal* **31**(7): 734-758

Hoang H, Rothaermel FT. 2005. The effect of general and partner-specific alliance experience on joint R&D project performance. *The Academy of Management Journal* **48**(2): 332-345

Hobday M. 2000. The project-based organisation: an ideal form for managing complex products and systems? *Research policy* **29**(7-8): 871-893

Hoegl M, Gemuenden HG. 2001. Teamwork quality and the success of innovative projects: A theoretical concept and empirical evidence. *Organization Science* **12**(4): 435-449

Hoegl M, Weinkauf K, Gemuenden HG. 2004. Interteam coordination, project commitment, and teamwork in multiteam R&D projects: A longitudinal study. *Organization Science* **15**(1): 38-55

Huckman RS, Zinner DE. 2008. Does focus improve operational performance? Lessons from the management of clinical trials. *Strategic Management Journal* **29**(2): 173-193

Jaffe AB, Trajtenberg M, Henderson R. 1993. Geographic localization of knowledge spillovers as evidenced by patent citations. *The Quarterly Journal of Economics*: 577-598

Jain S, George G. 2007. Technology transfer offices as institutional entrepreneurs: the case of Wisconsin Alumni Research Foundation and human embryonic stem cells. *Industrial and Corporate Change* **16**(4): 535

Jones BF, Wuchty S, Uzzi B. 2008. Multi-University Research Teams: Shifting Impact, Geography, and Stratification in Science. *Science* **322**(5905): 1259-1262

Kale P, Dyer JH, Singh H. 2002. Alliance capability, stock market response, and long-term alliance success: the role of the alliance function. *Strategic Management Journal*: 747-767

Kale P, Singh H. 2007. Building firm capabilities through learning: the role of the alliance learning process in alliance capability and firm level alliance success. *Strategic Management Journal* **28**(10): 981-1000

Katila R, Ahuja G. 2002. Something old, something new: A longitudinal study of search behavior and new product introduction. *Academy of Management Journal* **45**(6): 1183-1194

Katz JS, Martin BR. 1997. What is research collaboration? *Research Policy* **26**(1): 1-18

Katz R. 1982. The effects of group longevity on project communication and performance. *Administrative Science Quarterly*: 81-104

Kiesler S, Cummings JN. 2002. What do we know about proximity and distance in work groups? A legacy of research. *Distributed work* **1**: 57-80

Kraut R, Steinfield C, Chan AP, Butler B, Hoag A. 1999. Coordination and virtualization: The role of electronic networks and personal relationships. *Organization Science* **10**(6): 722-740

Kraut RE, Fussell SR, Brennan SE, Siegel J. 2002. Understanding effects of proximity on collaboration: Implications for technologies to support remote collaborative work. *Distributed work*: 137-162

Krishnan R, Martin X, Noorderhaven NG. 2006. When does trust matter to alliance performance? *Academy of Management Journal* **49**(5): 894-917

Kumar K, Van Fenema PC, Von Glinow MA. 2008. Offshoring and the global distribution of work: Implications for task interdependence theory and practice. *Journal of International Business Studies* **40**(4): 642-667

Kumar R, Nti KO. 1998. Differential learning and interaction in alliance dynamics: a process and outcome discrepancy model. *Organization Science*: 356-367

Lanjouw JO, Schankerman M. 2001. Characteristics of patent litigation: a window on competition. *RAND Journal of Economics*: 129-151

Lauring J. 2009. Managing cultural diversity and the process of knowledge sharing: A case from Denmark. *Scandinavian Journal of Management* **25**(4): 385-394

Lawrence PR, Lorsch JW. 1967. Differentiation and integration in complex organizations. *Administrative science quarterly* **12**(1): 1-47

Lerner J. 1995. Venture capitalists and the oversight of private firms. *Journal of Finance*: 301-318

Levitt B, March JG. 1988. Organizational learning. *Annual review of sociology* **14**(1): 319-338

Litwak E, Hylton LF. 1962. Interorganizational Analysis: A Hypothesis on Co-ordinating Agencies. *Administrative Science Quarterly* **6**(4): 395-420

Madhok A, Tallman SB. 1998. Resources, transactions and rents: managing value through interfirm collaborative relationships. *Organization Science*: 326-339

Makela K, Kalla HK, Piekkari R. 2007. Interpersonal similarity as a driver of knowledge sharing within multinational corporations. *International Business Review* **16**(1): 1-22

Malone TW, Crowston K. 1994. The interdisciplinary study of coordination. *ACM Computing Surveys (CSUR)* **26**(1): 87-119

March JG, Simon HA. 1958. *Organizations*. John Wiley & Sons Inc

McCormick J, Owen-Smith J, Scott C. 2009. Distribution of human embryonic stem cell lines: who, when, and where. *Cell stem cell* **4**(2): 107-110

Menard SW. 2002. *Applied logistic regression analysis*. Sage Publications Inc

Mendez A. 2003. The coordination of globalized R&D activities through project teams organization: an exploratory empirical study. *Journal of World Business* **38**(2): 96-109

Minbaeva D, Pedersen T, Bjorkman I, Fey CF, Park HJ. 2003. MNC knowledge transfer, subsidiary absorptive capacity, and HRM. *Journal of International Business Studies* **34**(6): 586-599

Miner AS, Bassoff P, Moorman C. 2001. Organizational improvisation and learning: A field study. *Administrative Science Quarterly*: 304-337

Murray F, O'Mahony S. 2007a. Exploring the Foundations of Cumulative Innovation: Implications for Organization Science. *Organization Science* **18**(6): 1006

Murray F, Stern S. 2007b. Do formal intellectual property rights hinder the free flow of scientific knowledge? An empirical test of the anti-commons hypothesis. *Journal of Economic Behavior and Organization* **63**(4): 648-687

Narula R. 2002. Innovation systems and 'inertia' in R&D location: Norwegian firms and the role of systemic lock-in. *Research Policy* **31**(5): 795-816

Narula R, Santangelo GD. 2009. Location, collocation and R&D alliances in the European ICT industry. *Research Policy* **38**(2): 393-403

Nelson RR, Winter SG. 1982. *An Evolutionary Theory of Economic Change*. Harvard University Press

Nerkar A. 2003. Old is gold? The value of temporal exploration in the creation of new knowledge. *Management Science* **49**(2): 211-229

Nooteboom B. 2000a. *Learning and innovation in organizations and economies*. Oxford University Press, USA

Nooteboom B. 2000b. Learning by interaction: absorptive capacity, cognitive distance and governance. *Journal of Management and Governance* **4**(1): 69-92

Nooteboom B, Van Haverbeke W, Duysters G, Gilsing V, van den Oord A. 2007. Optimal cognitive distance and absorptive capacity. *Research Policy* **36**(7): 1016-1034

O'Leary MB, Cummings JN. 2007. The Spatial, Temporal, and Configurational Characteristics of Geographic Dispersion in Teams.

Olson JS, Teasley S, Covi L, Olson G. 2002. The (currently) unique advantages of collocated work. *Distributed work*: 113-135

Orlikowski WJ. 2002. Knowing in practice: Enacting a collective capability in distributed organizing. *Organization Science*: 249-273

Owen-Smith J, Powell WW. 2003. The expanding role of university patenting in the life sciences: assessing the importance of experience and connectivity. *Research Policy* **32**(9): 1695-1711

Podolny JM, Stuart TE. 1995. A role-based ecology of technological change. *American Journal of Sociology*: 1224-1260

Polanyi M. 1967. The Tacit Dimension. *New York*

Powell WW, Koput KW, Smith-Doerr L, Owen-Smith J. 1999. Network Position and Firm Performance: Organizational Returns to Collaboration in the Biotechnology Industry. *Research in the Sociology of Organizations* **16**(1): 129-159

Powell WW, Koput KW, SmithDoerr L. 1996. Interorganizational collaboration and the locus of innovation: Networks of learning in biotechnology. *Administrative Science Quarterly* **41**(1): 116-145

Puranam P, Singh H, Chaudhuri S. 2009. Integrating Acquired Capabilities: When Structural Integration Is (Un) necessary. *Organization Science* **20**(2): 313

Reagans R, Argote L, Brooks D. 2005. Individual Experience and Experience Working Together: Predicting Learning Rates from Knowing Who Knows What and Knowing How to Work Together. *Management Science* **51**(6): 869

Rothaermel FT, Deeds DL. 2006. Alliance type, alliance experience and alliance management capability in high-technology ventures. *Journal of business venturing* **21**(4): 429-460

Rothaermel FT, Hess AM. 2007. Building dynamic capabilities: Innovation driven by individual-, firm-, and network-level effects. *Organization Science* **18**(6): 898-921

Rowley T, Behrens D, Krackhardt D. 2000. Redundant governance structures: An analysis of structural and relational embeddedness in the steel and semiconductor industries. *Strategic Management Journal*: 369-386

Sarbaugh-Thompson M, Feldman MS. 1998. Electronic Mail and Organizational Communication: Does Saying" Hi" Really Matter? *Organization Science* **9**(6): 685

Schreiner M, Kale P, Corsten D. 2009. What really is alliance management capability and how does it impact alliance outcomes and success? *Strategic Management Journal* **30**(13): 1395-1419

Schumpeter JA. 1939. *Business cycles*. McGraw-Hill New York

Scott C, McCormick J, Owen-Smith J. 2009. And then there were two: use of hESC lines. *Nature Biotechnology* **27**(8): 696-697

Shan W, Walker G, Kogut B. 1994. Interfirm cooperation and startup innovation in the biotechnology industry. *Strategic Management Journal* **15**(5): 387-394

Simon HA. 1997. *Administrative behavior: A study of decision-making processes in administrative organizations*. Free Press

Song J, Almeida P, Wu G. 2003. Learning-by-hiring: when is mobility more likely to facilitate interfirm knowledge transfer? *Management Science* **49**(4): 351-365

Sørensen JB, Stuart TE. 2000. Aging, obsolescence, and organizational innovation. *Administrative science quarterly* **45**(1): 81-112

Sorenson O, Fleming L. 2004. Science and the diffusion of knowledge. *Research Policy* **33**(10): 1615-1634

Srikanth K. 2010. Coordination in Distributed Organizations. *Working Paper Series, SSRN*: 1-37

Srikanth K, Puranam P. 2009. Coordination Within vs. Across Firm Boundaries. *Working Paper Series, SSRN*: 1-41

Taylor A, Greve HR. 2006. Superman or the fantastic four? Knowledge combination and experience in innovative teams. *Academy of Management Journal* **49**(4): 723-740

Tushman ML, Nadler DA. 1978. Information processing as an integrating concept in organizational design. *The Academy of Management Review* **3**(3): 613-624

Tzabbar D. 2009. When does scientist recruitment affect technological repositioning? *Academy of Management Journal* **52**(5): 873-896

Van de Ven AH, Polley D. 1992. Learning while innovating. *Organization Science* **3**(1): 92-116

Vanneste BS. 2009. The Organization of Vertical Relationships. 1-168

Watson WE, Kumar K, Michaelsen LK. 1993. Cultural diversity's impact on interaction process and performance: Comparing homogeneous and diverse task groups. *The Academy of Management Journal* **36**(3): 590-602

Weick K. 1997. Cosmos vs. Chaos: Sense and Nonsense in Electronic Contexts,". *Knowledge in organizations*

Weisband S. 2002. Maintaining awareness in distributed team collaboration: Implications for leadership and performance. *Distributed work*: 311–333

White S, Lui SSY. 2005. Distinguishing costs of cooperation and control in alliances. *Strategic Management Journal* **26**(10): 913-932

Wood RE. 1986. Task complexity: Definition of the construct. *Organizational behavior and human decision processes* **37**(1): 60-82

Wuchty S, Jones BF, Uzzi B. 2007. The Increasing Dominance of Teams in Production of Knowledge. *Science* **316**(5827): 1036

Yelle LE. 1979. The learning curve: Historical review and comprehensive survey. *Decision Sciences* **10**(2): 302-328

Zahra S, George G. 2002. Absorptive capacity: A review, reconceptualization, and extension. *Academy of Management Review* **27**(2): 185-203

Zollo M, Reuer JJ, Singh H. 2002. Interorganizational routines and performance in strategic alliances. *Organization Science* **13**(6): 701-713

Zucker L, Darby MR, Brewer MB, Peng Y. 1996. Collaboration Structure and Information Dilemmas in Biotechnology. *Trust in Organizations: Frontiers of Theory and Research*: 90-113

Appendix A

Interviews

Although I do not use as a primary data source in any of my studies, I have conducted several interviews to inform the theory behind my work and to shape my empirical approach. The first set of interviews that I have conducted is with scientists from the stem cell laboratories located in the United Kingdom. Having direct contact with scientists and getting their perspective on the day to day tasks related to both lab work and the discovery processes allowed me to more effective direct my research efforts and make better sense of my findings. In this regard, I have interviewed with two Principal Investigators as well as three lab researchers and a lab coordinator in stem cell research area. The principal investigators are not only prominent figures in the field but they also have patented inventions in the human embryonic stem cell research and show up in my dataset, hence their insights were greatly valuable in making sense of my results. The second set of interviews that I have conducted is with representatives of two of the prominent clinical trial design and data management organizations, RapidTrials Inc. and MediData Solutions Ltd. These organizations work extensively with both trial sponsors and trial sites and help with the design, management and execution of trials. As such, their insights were highly valuable in helping me understand the clinical trial process.

In my interview with Professor Harry Moore, he has pointed to the detailed importance of background in the stem cell research area and how that background helped access to cell lines. The main point that resonated was the existence of prior information on the subject of stem cells. Although my particular setting is the human embryonic stem cell domain (hESC), this domain does have extensive links to the broader area of stem cell research. Prof. Moore highlighted

several factors; first that the fundamental characteristics of hESC, allow for wide variety of participant to get interested in the area. Second, that despite the initial aura in this area, regeneration of tissue is quite complex in both proliferation and maintenance. Third, the assays and techniques required for maintenance of the new cell lines is complex and although simple tests are easier conducted, research leading to tissue generation has multiple facets in terms of structure and cohesion of cells that is beyond simple cell regeneration.

During my interview with Professor Peter Andrews, Prof. Andrews shared his view on the evolution of research in human embryonic stem cells and gave anecdotes as to how he first got interested and started working in this area. He mentioned that he gained access to the first cell lines in hESC through his personal contact with Professor James Thomson of Wisconsin University. When I asked about the extent of interdepartmental collaborations and if there were any problems in coordinating research across departments, he said that his research partners are usually from within the university and in cases where collaborations occur across departments they would be using on-site meetings rather than other communication channels for coordinating action. He said that this is because in his opinion the best observations occur in the actual lab and it's easiest to understand the progress of research and resolve problems on site although it would be possible share results easily otherwise as well. Prof. Andrews also pointed to the fact that although basic steps in research can be and are easily replicable, the use of multiple approaches necessary to integrate the knowledge during a research project in hESC is quite complex especially in the case of tissue regeneration. He concluded that most collaborations that are across departments focus on more of the basic questions regarding hESC mechanics such as cell replication and differentiation rather than tissue or organ regeneration and that this is potentially

due to the early stage of the research and also due to the high complexity involved in alternative projects like tissue generation.

I have also had the chance to get insights from a PhD student at Professor Andrews' lab, he gave detailed information as to how cell proliferation takes place and what basic questions are addressed in the current state of art of hESC research. He further gave an explanation as to how he came into contact with Professor Andrews and decided to join the lab. This information was helpful in understanding how patterns of collaboration evolve in this area; in his case, it was through an initiation on his part while he was a researcher in China. More importantly, it also shows that even though he has dual affiliations in both China and the UK, as it shows up on his published work, his main line of work is taking place in the UK lab. This proves that one needs to be cautious of the disparity between disclosed affiliation and actual work location and that the importance of personal contacts is especially important in initiating projects in this line of work.

At the same time, I had the chance to talk with Mrs. Angela Ford who is coordinating the stem cell lab at Sheffield University. She explained the initiatives taken at the academic level to bring together stem cell scientists in the UK. According to her, even though the stem cell community is quite small and the scientists often exchange ideas through conferences and meetings, the bulk of research in the area still takes place in the individual labs. In instances where there is collaboration across institutions, the collaboration is enhanced through exchanges of students and through visiting faculty. The funding for such collaborations usually comes from beyond the university through support program such as the ESTOOLS (European Commission Framework Programme) and the like.

In addition, I was able to have brief interviews with researchers from University College London Centre for Stem Cell and Regenerative Medicine and from the Stem Cell Biology Group

at Imperial College London. The views of researchers from the two institutions converged with those of the scientists from previous interviews. Although the basic steps in replicating cell lines can be easy, the actual discovery process is complex and teams are usually formed to work together in one particular lab. Exceptions do occur, however, usually are accommodated with some sort of coordination mechanism involving collocation or frequent visits.

I have also had interviews with representatives of RapidTrials Inc. and MediData Solutions Ltd. companies specializing in clinical trial design and data management. The aim of these interviews was twofold; a)to get a clear understanding of the way in which clinical trials are designed, in terms of selection of tasks and trial sites and b)to understand how the involvement of clinical research organizations effect the management of clinical trials. Although I didn't have the opportunity to access the databases of these two companies upon my interviews, my conversations did give me a clearer picture of the field and the way in which clinical trials are organized. In terms of the clinical research organizations' involvement in clinical trials, I have decided to construct a measure to control for clinical research organization involvement through the use of a business directory (Buyer's Directory) which lists all clinical research organizations by year.

Overall, my interviews were essential in helping me understand how the two collaborations work in the two contexts of my interest. The valuable insights and information that I received through my interviewees helped me define my theoretical angle and construct the right measures for my empirical approaches. Although I do not use these interviews as a primary data source, having conducted these interviews was critical for the development of my studies especially in Chapters 2 and 4.